REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0762

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December, 1997	

4. TITLE AND SUBTITLE
1997 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 7B, Armstrong Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Armstrong Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1997

GRADUATE STUDENT RESEARCH PROGRAM
FINAL REPORTS

VOLUME 7B

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1997

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 7A is bound in two parts, 7A and 7B. Volume 7A contains #1-14. Volume 7B contains reports #15-29. The Table of Contents for Volume 7 is included in both parts.

This document is one of a set of 16 volumes describing the 1997 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4A & 4B	Rome Laboratory
5A, 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, U.S. Air Force Academy, Air Logistic Centers, and Wilford Hall Medical Center
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, U.S. Air Force Academy, Air Logistic Centers, and Wilford Hall Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the

1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435
1997	148	98	140	427

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1997 are shown in this table.

1997 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	490	188	32
(HBCU/MI)	(0)	(0)	(0)
GSRP	202	98	9
(HBCU/MI)	(0)	(0)	(0)
HSAP	433	140	14
TOTAL	1125	426	55

4. SITE VISITS

During June and July of 1997, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1997 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1997 SRP selected participants are shown here.

1997 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	89	123
USAF Laboratory Funds	48	9	17
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	0	0	N/A
TOTAL	9	98	140

SFRP - 188 were selected, but thirty two canceled too late to be replaced.
GSRP - 98 were selected, but nine canceled too late to be replaced.
HSAP - 140 were selected, but fourteen canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1997 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996	1997
Faculty Members	$690	$718	$740	$740	$740	$770	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % (194 out of 286) claimed travel reimbursements at an average round-trip cost of $776.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of $388. By contrast, in 1993, 58 % of SFRP associates took

orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the second year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, only 7% of the associates' information needed clarification in order for direct deposit to properly function as opposed to 10% from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1997 REPORT

The complete set of reports for the 1997 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1997 associates, as indicated below:

1997 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions. GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	189
GSRP	97
HSAP	140
TOTAL	426

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	0	184
Master's	2	41	43
Bachelor's	0	56	56
TOTAL	186	97	298

SFRP Academic Titles	
Assistant Professor	64
Associate Professor	70
Professor	40
Instructor	0
Chairman	1
Visiting Professor	1
Visiting Assoc. Prof.	1
Research Associate	9
TOTAL	186

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1997 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
USAFA	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
   a. Time available for evaluation of applications:
   b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5. a. Benefits of associate's work to laboratory:
   b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
   b. Enhancement of research qualifications for SFRP associate:
   c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
   b. Enhancement of knowledge for SFRP associate:
   c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*USAFA*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1997 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1997 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1997 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1997 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

A PROTOCOL FOR DEVELOPMENT OF AMPLICONS FOR A RAPID AND EFFICIENT METHOD OF GENOTYPING HEPATITIS C VIRUS FROM CLINICAL SERUM SPECIMENS

Thomas R. Mertz Jr.
Graduate Student
Institute of Molecular Biology and Medicine
University of Scranton
Scranton, PA 18510

and

John R. Herbold
Associate Professor
School of Public Health
UT-Houston Health Science Center
Houston, TX 77030

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC

and

Armstrong Laboratory

August 1997

A PROTOCOL FOR DEVELOPMENT OF AMPLICONS FOR A RAPID AND EFFICIENT METHOD OF GENOTYPING HEPATITIS C VIRUS FROM CLINICAL SERUM SPECIMENS

Thomas R. Mertz Jr.
Graduate Student
Institute of Molecular Biology and Medicine
University of Scranton
Scranton, PA 18510
and
John R. Herbold
Associate Professor
School of Public Health
UT-Houston Health Science Center
Houston, TX 77030

Abstract

The two goals of this research project were to (1) establish a procedure to confirm the presence of Hepatitis C Virus (HCV) in serum and (2) to establish a procedure to obtain unique amplicons suitable for genotyping HCV specimens. HCV detection needed to be simple and rapid for use in a clinical laboratory setting. Genotyping of HCV is necessary for further epidemiological studies and assessment of risk to military personnel deployed worldwide. This research examines two protocols for (1) a PCR-based assay to test for HCV RNA (antigen) and (2) isolating viral RNA and preparing cDNA for a PCR-based method for genetic typing of HCV variants. Our results indicate the potential for exploring novel applications of current technology to rapidly genotype HCV variants in the clinical laboratory setting. This work will continue throughout the year at the Institute of Molecular Biology and Medicine (University of Scranton) in collaboration with the Air Force Epidemiology Research Division (Armstrong Laboratory) and the University of Texas School of Public Health.

A PROTOCOL FOR DEVELOPMENT OF AMPLICONS FOR A RAPID AND EFFICIENT METHOD OF GENOTYPING HEPATITIS C VIRUS FROM CLINICAL SERUM SPECIMENS

Thomas R. Mertz Jr. and John R. Herbold

INTRODUCTION

Hepatitis C Virus (HCV) is a single stranded RNA virus belonging to the Flaviviridae family. It is known to be the major causative agent of non-A, non-B hepatitis(Weiner et al., 1990; Choo et al., 1991). Infection can be asymptomatic, mild, chronic, and/or result in long term sequelae such as a chronic viral carrier status and potentially liver cancer. Individuals acutely infected with HCV cannot readily be differentiated on clinical or histopathologic criteria from those infected with other viral hepatitides, and, in many cases, may show no clinical manifestations. Antibody formation may not occur until several months after initial infection(Young et al., 1993). Parenteral and sexual transmission are believed to be the primary routes of infection(Houghton, 1996). Infectivity, pathogenicity, chronicity, and response to therapeutic intervention may be related to the specific HCV genotype causing infection(Hagiwara et al., 1996; Shindo et al., 1991; Versalovic et al., 1996; Yuki et al., 1997). But at present, no means is generally available to establish genotype.

Six genotypes as well as several sub-types have been elucidated after comparison of specific sequences(Parsing and Relman, 1996; Simmonds, Alberti et al., 1994; Simmonds, Holmes et al., 1993). Most work to date has focused on two distinct regions of the HCV genome: the 5' non-coding region due to its high degree of conservation, and the NS-5 region due to its variability (Cha et al., 1991; Bukh et al., 1992b; Choo et al., 1991; O'Brien et al., 1997; Ohno et al., 1997; Okamoto et al., 1992; Pernas, 1995; Simmonds, Smith et al., 1994; Simmonds, McOmish et al., 1993). Because PCR primer selection was very important in this study, and the 5'-NCR and NS-5 region are commonly used for genotypic work, we chose one primer set from each region using the Oligo 5.0 software package for Macintosh (Rychlik, 1996).

Reverse transcription was carried out with random primers, and thus the entire genome was amplified into DNA. This result was beneficial because it provides nucleic acid for future

analysis of different areas of the genome. Also, a DNA form of the genome is more stable in storage than the RNA form.

METHODOLOGY

I. Current Clinical Laboratory Diagnostic Protocol.

Archived samples were obtained from the Air Force Epidemiological Reference Laboratory. Clinical serum samples are screened for hepatitis C antibody at the primary military medical facility by an Enzyme-Linked Immunosorbent Assay (ELISA) procedure (Abbott, 1993). ELISA positive serum samples are forwarded to the Epidemiologic Research Division for confirmation of results by repeat ELISA and a Strip Immunoblot Assay (RIBA)(Chiron, 1995; Chaudhary et al., 1993). Antibody results are provided to the attending physician for assessment in conjunction with other laboratory parameters and the clinical status of the patient.

II. Implementation of Hepatitis C Viral RNA (Antigen) Test Procedure.

Our goal was to integrate an HCV antigen test procedure into the daily activities of a clinical reference laboratory so that only antigen positive samples would be subjected to HCV viral RNA isolation procedures. Samples used for antigen screening were all repeatedly reactive by ELISA and had variable (indeterminate and positive) RIBA results.

Sample sera were tested for HCV antigen using the AMPLICOR HCV Test Kit(Roche, 1993). Approval was received from Roche Molecular Systems to participate in the RMS Research Use Certification Program for the AMPLICOR HCV Test Kit. The AMPLICOR HCV Test Kit has not been approved by the FDA, and, as such, is not used for diagnostic purposes. Specimens were prepared and tested as stated on the product insert. A 100 microliter aliquot of serum was required for each test. All necessary reagents for reverse transcription of viral RNA, amplification by PCR, and detection of the amplicon were included in the kit. Laboratory technicians and AFOSR personnel were trained by a Roche representative in the use of the kit and good laboratory practices for PCR procedures.

III. Isolation of Hepatitis C Viral RNA.

Isolation of viral RNA from sera was accomplished using a commercially available kit (QIAmp Viral RNA Kit, QIAGEN Inc. #29504) in accordance with the handbook procedures. A 140 microliter aliquot of serum was required for each test. The concentration of purified RNA in each sample was determined spectrophotometrically(see Table 1) (Pharmacia Biotech Ultrospec 3000 UV/Vis Spectrophotometer). Purified RNA was either immediately reverse-transcribed or stored at -20 C until ready for use (see Methodology Section IV, Generation of Amplicons for Genotyping).

IV. Generation of Amplicons for Genotyping .

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) was accomplished using a commercially available kit (RT-PCR Kit, Stratagene Inc., #200420). RT-PCR is used to generate amplicons of complementary DNA from RNA. Random primers were used in the synthesis of cDNA. The use of random primers allowed the formation of amplicons that, collectively, represent the entire genome. Further PCR amplification with sequence-specific primers(see Methodology Section V, Selection of Sequence-Specific Primers for Genotyping) generated amplicons suitable for sequence analysis and genotyping(Stratagene, 1997). Reverse transcriptase synthesis of first-strand cDNA requires 100 ng of RNA, according to the product insert. Volumes required ranged from 1.1-16.0 microliters based on the concentration of the previously purified RNA (Table 1).

V. Selection of Sequence-Specific Primers for Genotyping.

Specific sections in the 5'NCR and NS-5 regions were targeted based on a literature review indicating that there are two regions of the genome useful for phylogenetic analysis (Chan et al., 1992; Okamoto et al., 1996; Simmonds et al., 1994). Any sequence or primer data is in reference to Genbank accession number L02836. Potential primers (Table 2) for amplification of the NS-5 and 5'NC regions were analyzed by Oligo 5.0 for Macintosh (Rychlik, 1996). Primer selections were determined by best fit based on factors such as primer Tm matching, GC clamp

design, optimal annealing temperatures, and dimer formation potential. Primers chosen for subsequent PCR are also noted in Table 2.

RESULTS AND DISCUSSION

I. Constraints of Current Clinical Laboratory Diagnostic Protocol.

Requests for assessment of Hepatitis C viral antigen status have been received but there are currently no FDA approved diagnostic kits available. Additionally, research studies have indicated a variety of hepatitis C genotypes that appears to be related to infectivity, pathogenicity, chronicity, and response to therapeutic intervention. Assessment of hepatitis C viral antigen status is useful for multiple purposes. One, viral presence differentiates the active hepatitis case from patients who have been infected but are now recovered. Also, viral load is helpful when assessing response to antiviral therapy. Thirdly, viral presence coupled with epidemiologic characteristics may elucidate the natural history of hepatitis C virus infections leading to clues regarding prevention, intervention, and therapeutic treatment. Genotyping (currently a basic research procedure) of the different viral isolates would facilitate identification and elucidation of epidemiologic patterns related to infectivity, pathogenicity, chronicity, as well as response to therapeutic intervention(Enomoto et al., 1990; Murashima et al., 1996; Murphy et al., 1994; Nagasaka et al., 1996; Matar et al., 1996; McOmish et al., 1994).

II. Implementation of Hepatitis C Viral RNA (Antigen) Test Procedure

The Roche Molecular Systems AMPLICOR HCV Test Kit was easily integrated into the daily activities of a clinical reference laboratory. It was found that the test could easily be completed within the course of a normal work day, with results usually available in 5-8 hours depending on sample size. Once FDA approval is obtained this procedure could easily be added to existing clinical reference laboratory protocols.

III. Isolation of HCV viral RNA

The QIAmp kit was chosen due to its ease of use and processing speed. The kit recovers viral RNA by passing a previously lysed sample through a column that selectively binds viral RNA. A series of washes eliminates contaminants and elutes purified RNA in water (Qiagen, 1996). Viral RNA was isolated from 50 serum samples. Purified RNA concentrations are displayed in Table 1. Some difficulty was experienced in obtaining significant quantities of purified RNA (Table 1). In order to insure greater yields, larger aliquots of serum can be processed using the QIAmp Viral RNA Kit. If greater sensitivity is desired due to low viral titer in the serum or if serum samples have undergone repeat freeze thaw cycles, it may be necessary to increase the volume of material processed through the QIAmp spin column. The procedure is outlined in the Protocol for Large Volume Samples (Qiagen, 1996). Although the handbook suggests that the viral RNA isolation can be completed in approximately 20 minutes, we found that the time for completion depended greatly upon sample size. Thus, times for processing ten samples averaged about one hour.

IV. Generation of Amplicons for Genotyping.

RT was accomplished initially on 13 purified viral RNA samples. Samples were forwarded to the Institute of Molecular Biology and Medicine for evaluation of the primer set selected for genotyping. In some instances, it appeared that RT of purified RNA was unsuccessful. Therefore, the protocol for Synthesis of First-Strand cDNA using RT was changed as follows:

(1) Independent of the concentration of purified RNA, a 10 microliter aliquot of RNA suspension was used (RNA quantities ranged from 250-1000 ng).

(2) A master mix of buffer, Rnase Block Ribonuclease Inhibitor, dNTPs, and MMLV-RT was prepared and added in aliquots to the reactions following the 10 minute cooling period that allowed the primers to anneal to the RNA.

V. Selection of Specific Primers for Genotyping.

If the sequence data provides significant genotypic differentiation, research will be conducted into the ability to determine genotype by analysis of the amplicon melting curve during PCR (Ririe et al., 1997). This would extend greatly the speed at which genotype can be determined and practicality in a clinical reference laboratory setting.

CONCLUSION

Integration of currently available clinical test kits and standard research procedures can greatly enhance the information obtained relative to HCV infection of individuals and their clinical prognosis. The Roche HCV RNA procedure directs that a specimen diluent consisting of a bicine buffered solution containing manganese, potassium, and 0.05% sodium azide be used to resuspend the RNA pellet. The Stratagene RT procedure requires the use of RNA previously suspended in DEPC water. If the Roche test procedure diluent is compatible with the Stratagene RT procedure, or if the Roche RNA pellet could be resuspended in DEPC treated water, then the reisolation of viral RNA step could be eliminated, adding to the efficiency of the procedure. Technical Services representatives of both kit manufacturers were unable to address this issue. Parallel tests should be run to determine compatibility of the diluents with either kit for research purposes. Modification of the protocol as described resulted in generation of adequate cDNA for storage and/or transport to offsite locations for additional research procedures such as gene sequencing. The amplicons produced will be used in future sequence studies at the University of Scranton Institute of Molecular Biology and Medicine in collaboration with the Armstrong Laboratory and the University of Texas School of Public Health. If the sequence data provides significant genotypic differentiation, research will be conducted into the ability to determine

genotype by analysis of the amplicon melting curve during PCR(Ririe et al., 1997). This would extend greatly the speed at which specific genotypes can be determined.

The availability of archived HCV positive sera from a broad spectrum of military personnel worldwide provides an opportunity to investigate the association between HCV genotype and risk of infection - an issue of significant military importance. The potential for widespread HCV infection in U.S. military populations parallels the issues raised when HIV was first identified (Herbold, 1986). HCV infection can be transmitted parenterally (blood transfusion), may be transmitted sexually, can result in chronic impairment, and can result in a infectious carrier state.

Elucidation of infection patterns in U.S. military populations worldwide will depend on the identification of the predominant variants at the gene sequence level (Smith et al., 1997). This information coupled with analysis of occupational, demographic, and behavioral risk factors will enable military planners to implement effective prevention and intervention programs concomitant with the level of risk encountered (McDade & Anderson, 1996). Additionally, this information will provide valuable insight as to whether a total force HCV screening program should be implemented for this emerging infectious disease risk.

Table 1

Sample		Conc. (ng/ul)		ug used	
c1		58.0		1.7241	
c2		69.0		1.4493	
c3		41.5		2.4096	
cr4		57.0		1.7544	
c5		63.0		1.5873	
c6		58.4		1.7123	
c7		34.6		2.8902	
c8		55.0		1.8182	
c9		57.2		1.7483	
c10		60.6		1.6502	
c11		98.6		1.0142	
c12		29.1		3.4364	
c13		93.6		1.0684	
c14		67.0		1.4925	
c15		78.5		1.2739	
c16		53.5		1.8692	
c17		25.5		3.9216	
c18		26.8		3.7313	
c19		53.7		1.8622	
c20		75.6		1.3228	
c21		88.0		1.1364	
c22		20.9		4.7847	
c23		75.3		1.328	
c24		76.1		1.3141	
c25		11.6		8.6207	
c26		63.3		1.5798	
c27		48.9		2.045	
c28		79.7		1.2547	
c29		86.3		1.1587	
c30		11.3		8.8496	
c31		82.7		1.2092	
c32		32.8		3.0488	
c33		46.4		2.1552	
c34		0		0.000000	
c35		76.1		1.3141	
c36		39.7		2.5189	
c37		72.9		1.3717	
c38		13.2		7.5758	
c39		29.1		3.4364	
c40		11.4		8.7719	
c41		11.4		8.7719	
c42		56.4		1.773	
c43		21.1		4.7393	
c44		61.1		1.6367	
c45		60.7		1.6474	
c46		7.5		13.333	
c47		28.6		3.4965	
c48		6		16.667	
c49		29.3		3.413	
c50		27.1		3.69	

Table 2

Name	Region	5' base	Polarity	5'-3' Sequence	Reference
242	NS-5	8304	-	GGCGGAATTCCTGGTCATAGCCTCCGTGAA	Chan et al, 1992
555	NS-5	8227	-	CCACGACTAGATCATCTCCG	Chan et al, 1992
243	NS-5	7904	+	TGGGGATCCCGTATGATACCCGCTGCTTTGA	Chan et al, 1992
554	NS-5	7935	+	CTCAACCGTCACTGAACAGGACAT	Chan et al, 1992
475	5'-NCR	319	+	GGAGGTCTCGTAGACCGTGC*	Okamoto et al. 1996
186	5'-NCR	751	-	ATGTACCCCATGAGGTCGGC*	Okamoto et al. 1996
O-1	NS-5	8275	+	CAGAACGACATCCGTGTT*	Oligo v 5.0 for Mac
O-2	NS-5	8769	-	GGAGTTGACTGGAGTGTGTC*	Oligo v 5.0 for Mac
				* Chosen for PCR	

REFERENCES

Abbott Laboratories. (1993). Abbott HCV EIA 2.0.

Bukh, J., Purcell, R., & Miller, R. (1992a). Importance of primer selection for the detection of hepatitis C virus RNA with the polymerase chain reaction assay. Proc Natl Acad Sci, 89:187-191.

Bukh, J., Purcell, R., & Miller, R. (1992b). Sequence analysis of the 5' noncoding region of hepatitis C virus. Biochemistry, 89:4942-4946.

Cha, T., Kolberg, J., & Irvine, B.e. (1991). Use of a Signature Nucleotide Sequence of Hepatitis C Virus for Detection of Viral RNA in Human Serum and Plasma. J Clin Micro, 29:(11). 2528-2534.

Chan, S.W., McOmish, F., & Holmes, E.C.e. (1992). Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. J Gen Virol, 73:1131-1141.

Chaudhray, R., Andonov, A., & MacLean C. (1993). Detection of Hepatitis C Virus Infection With Recombinant Immunoblot Assay, Synthetic Immunoblot Assay, and Polymerase Chain Reaction. J Clinical Laboratory Analysis, 7:164-167.

Chiron Corporation. (1995). Chiron RIBA HCV 2.0 SIA. 10/95 ed.

Choo, Q.L., Weiner, A.J., Overby, L.R., & et al. (1990). Hepatitis C virus: The major causative agent of viral non-A, non-B hepatitis. British Medical Bulletin, 46:(2). 423-441.

Choo, Q., Richman, K.H., Han, J., & et al. (1991). Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci, 88:2451-2455.

Enomoto, N., Takada, A., Nakao, T., & Date, T. (1990). There are two major types of the hepatitis c virus in Japan. Biochemical and Biophysical Research Communications, 170:1021-1025.

Hagiwara, H., Hayashi, N., Kasahara, A., & et al. (1996). Treatment with Recombinant Interferon-alpha2a for Patients with Chronic Hepatitis C: Predictive factors for Biochemical and Virologic Response. Scand J Gastroenterol, 31:1021-1026.

Herbold Jr. (1986). AIDS Policy Development Within the Department of Defense. Military Medicine, 151:(12). 623-627.

Houghton, M. (1996). Hepatitis C Viruses. In B.N. Fields, D.M. Knipe, P.M. Howley, & et al. (Eds.), Fields Virology. (pp. 1035-1051). Philadelphia: Lippincott-Raven.

Matar, G., Sharrara, H., Abdelnour, G., & Abdelnoor, A. (1996). Genotyping of Hepatitis C Virus Isolates from Lebanese Hemodialysis Patients by reverse Transcription-PCR and Restriction Fragment Length Polymorphism Analysis of 5' Noncoding Region. J Clin Micro, 34:(10). 2623-2624.

McDade, J.E., & Anderson, B.E. (1996). Molecular Epidemiology: Applications of Nucleic Acid Amplification and Sequence Analysis. Epidemiologic reviews, 18:(1). 90-97.

McOmish, F., Yap, P., Dow, B., & et al. (1994). Geographical Distribution of Hepatitis C Virus genotypes in Blood Donors: and International Collaborative Survey. J Clin Micro, 32:(4). 884-892.

Murashima, S., Sata, M., Suzuki, H., & et al. (1996). Sequence Variation of the Hypervariable Region in HCV Carriers with Normal ALT Levels: A Comparison with Symptomatic Carriers. Microbiol Immunol, 40:(12). 941-947.

Murphy, D., B.Willens. (1994). Use of the Non-Coding Region for the Genotyping of Hepatitis C Virus. J Infect Dis, 169, 473-474.

Nagasaka, A., Hige, S., Tsunematsu, I, & et al. (1996). Changes in Hepatitis C Virus Quasispecies and density Populations in Patients Before and After Interferon Therapy. J Medical Virology, 50:214-220.

O'Brien, C., Henzel, B., Wolfe, L., & et al. (1997). cDNA Sequencing of the 5' Noncoding Region (5' NCR) to Determine Hepatitis C Genotypes in Patients with Chronic Hepatitis C. Digestive Diseases and Sciences, 42:(5). 1087-1093.

Ohno, T., & et al. (1997). New Hepatitis C Virus (HCV) Genotyping System That Allows for Identification of HCV Genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J Clin Micro, 34:(1). 201-207.

Okamoto, H., & Kurai, K.e. (1992). Full-Length Sequennce of a Hepatitis C Virus Genome having poor homology to reported isolates: comparative dtudy of four distant genotypes. Virology, 1, 331-341.

Pernas, M., Bartolome, J., Castillo, I., & et al. (1995). Sequence of non-structural regions 3 and 5 of heptitis C virus genomes from Spanish patients: existence of a predominant variant related to type 1b. J Gen Virol, 76:415-420.

Persing, D.H., & Relman, D.A. (1996). Genotypic Detection of Antimicrobial Resistance. In D.H. Persing (Ed.), PCR Protocols for Emerging Infectious Diseases. (pp. 46-70). Washington, D.C. ASM Press.

Qiagen Inc. (1996). QIAmp Viral RNA Handbook, 6/96.

Ririe, K.M., Rasmussen, R.P., & Wittiwer, C.T. (1997). Product Differientation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. Analytical Biochemistry, 245:154-160.

Roche. (1993). Amplicor PCR Diagostics Hepatitis C Virus (HCV) Test. Instruction packet for test kit.

Rychlik, W. (1996). Oligo Primer Analysis Software v 5.0 for Mac. NBI/Genovus.

Shindo, M., Di Bisceglie, A.M., Cheung, C., & et al. (1991). Decrease in Serum Hepatitis C Viral RNA during Alpha-Interferon Therapy for Chronic Hepatitis C. Annals of Internal Medicine, 115:700-704.

Simmonds, P., Alberti, A., Alter, H.J., & et al. Editor & Hepatology 19:1321-1324. (1994). A Proposed System for the Nomenclature of Hepatitis C Viral Genotypes.

Simmonds, P., Holmes, E.C., Cha, T.A., & et al. (1993). Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J Gen Virol, 74:2391-2399.

Simmonds, P., McOmish, F., Yap, P.L., & et al. (1993). Sequence variability in the 5' non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. J Gen Virol, 74:661-668.

Simmonds, P., Smith, D.B., McOmish, F., & et al. (1994). Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. J Gen Virol, 75:1053-1061.
Smith, D.B., Pathirana, S., Davidson, F., & et al. (1997). The origin of hepatitis C genotypes. J Gen Virol, 78:321-328.

Stratagene Inc. (1997). Stratagene RT-PCR Kit Instruction Manual. 17005th ed.

Versalovic, J., Swanson, D.S., & Musser, J.M. (1996). Nucleic Acid Sequencing Studies of Microbial Pathogens: Insights into Epidemiology, Virulence, Drug Resistance, and Diversity. In D.H. Persing (Ed.), PCR Protocols for Emerging Infectious Diseases. (pp. 71-86). Washington, D.C: ASM Press.

Weiner, A.J., Kuo, G., Bradley, D.W., & et al. (1990). Detection of hepatitis C viral sequences in non-a, non-b hepatitis. Lancet,335:1-3.

Young, K.K.Y., Resnick, R.M., & Myers, T.W. (1993). Detection of Hepatitis C Virus RNA by a Combined Reverse Transcription-Polymerase Chain Reaction Assay. J Clin Micro, 31:(4). 882-886.

Yuki, N., Hayashi, N., Moribe, T., & et al. (1997). Relation of Disease Activity During Chronic Hepatitis C Infection to Complexity of Hypervariable Region 1 Quasispecies. Hepatology, 25:(2). 439-444.

AN PSYCHOMETRIC EXAMINATION OF THE MULTIDIMENSIONAL WORK ETHIC PROFILE AMONG AIR FORCE ENLISTED PERSONNEL

Michael J. Miller
Graduate Student
and
David J. Woehr
Associate Professor

Department of Psychology
Texas A&M University
College Station, Texas 77843-4235

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory
Human Resource Directorate
Brooks Air Force Base, TX

August 1997

AN PSYCHOMETRIC EXAMINATION OF THE MULTIDIMENSIONAL WORK ETHIC PROFILE AMONG AIR FORCE ENLISTED PERSONNEL

David J. Woehr
Associate Professor
and
Michael J. Miller
Graduate Student

Department of Psychology
Texas A&M University
College Station, Texas 77843-4235

Abstract

The present study examines the psychometric properties of the Multidimensional Work Ethic Profile (MWEP) developed by Michael Miller and David Woehr (Woehr & Miller, 1997, Miller and Woehr, 1997) with Air Force enlisted personnel. The MWEP is a multidimensional measure of work ethic based on previous literature and research focusing on work ethic and job performance. Originally developed based on a sample of university students, the MWEP has demonstrated good psychometric characteristics including reliability and validity. The MWEP has been suggested as a potentially valuable screening tool with Air Force enlisted personnel. The purpose of the present study was to provide a preliminary evaluation of the measure among Air Force enlisted personnel. Results indicate that the measure does demonstrate similar psychometric characteristics among Air Force enlisted personnel as with the original developmental sample. The MWEP provides reliable and valid measures of multiple dimensions underlying the work ethic construct. These results indicate that the MWEP may be a useful screening tool for Air Force Personnel.

AN PSYCHOMETRIC EXAMINATION OF THE MULTIDIMENSIONAL WORK ETHIC PROFILE AMONG AIR FORCE ENLISTED PERSONNEL

David J. Woehr
and
Michael J. Miller
Texas A&M University

Introduction

History and Definition of Work Ethic

The term "work ethic" was coined centuries ago by post-Reformation intellectuals who opposed the practice of social welfare and professed the importance of individualism (Byrne, 1990). They espoused the belief that human beings must assume full responsibility for their lot in life and the poor were no exception. As such, hard work was viewed as a panacea and through it, one could improve his or her condition in life. Implicit in this assumption was the belief that the poor simply needed to help themselves through diligent labor and all life's ills would vanish. Such were the harsh origins of the construct.

Modern formulations of the work ethic construct stem from the work of the German scholar Max Weber. It was in 1904 and 1905 that Weber wrote a two-part essay entitled "The Protestant Ethic and the Spirit of Capitalism". In this essay Weber advanced the thesis that the introduction and rapid expansion of capitalism and the resulting industrialization in Western Europe and North America was in part the result of the Puritan value of asceticism (i.e., scrupulous use of time, strict self-denial of luxury, worldly pleasure, ease, and so on to achieve personal discipline) and the belief in a calling from God (Byrne, 1990; Charlton, Mallinson, & Oakeshott, 1986; Fine, 1983; Furnham, 1990a; Green, 1968; Lehmann, 1993; Maccoby, 1983; Nord, Brief, Atieh, & Doherty, 1988; Poggi, 1983). It was the practice of asceticism that Weber believed produced the celebrated 'work ethic'--the complete and relentless devotion to one's economic role on earth (Lessnoff, 1994). An individual's economic role was prescribed by the belief in a calling (Gilbert, 1977). The manifestation of occupational rewards through success in one's calling came to be revered as a sign of being one of the elect (i.e., chosen by God to receive salvation). Thus, economic activity was a vehicle toward economic success and economic success was a sign of salvation.

Weber maintained that other Protestant faiths (e.g., Calvinism, Methodism, Pietism, and Baptists) shared common theological underpinnings in terms of being proponents of asceticism and the spirit of capitalism (Bouma, 1973; Nelson, 1973); thus the term "Protestant Work Ethic" (PWE). However, the premise that work ethic is a religiously oriented concept was contested then and since. In fact, researchers have found little relationship between religious orientation and endorsement of the work ethic (Giorgi & Marsh, 1990; Ray, 1982). Ray (1982) concluded that all religious orientations currently share the attributes associated with the work ethic to the same degree. He states that the Protestant ethic, "...is certainly not yet dead; it is just no longer Protestant" (p. 135). This is consistent with Pascarella's (1984) contention that all major religions have espoused the importance of work. Thus, it appears that what was originally conceived as a religious construct is now likely secular and is best viewed as general work ethic and not the PWE.

Since work ethic is not a surrogate for religious orientation the question becomes, What is it? Current conceptualizations tend to view work ethic as an attitudinal construct pertaining to work oriented values. An individual espousing a high work ethic would place great value on: hard work, autonomy, fairness, wise and efficient use of time, delay of gratification, and the intrinsic value of work (Cherrington, 1980; Dubin, 1963; Furnham, 1984; Ho & Lloyd, 1984; Weber, 1958; Wollack, Goodale, Wijting, & Smith, 1971). Therefore, work ethic seems to be made up of multiple components. These components appear to include: industriousness, asceticism, self-reliance, morality, delay of gratification, and the centrality of work. In the absence of a firmly accepted conceptual and operational definition it is posited that work ethic is a construct that reflects a constellation of attitudes and beliefs pertaining to work oriented behavior. Characteristics of "work ethic" are that it: (a) is multidimensional; (b) pertains to work and work related activity in general, not specific to any particular job (yet may generalize to domains other than work - school, hobbies, etc.); (c) is learned (not dispositional); (d) refers to attitudes and beliefs (not necessarily behavior); (e) is intendend as a motivational construct (should be reflected in behavior); and (e) is secular, not necessarily tied to any one set of religious beliefs.

Relevance of Work Ethic to the Air Force

As previously defined, individual differences in work ethic should reflect differences among individuals in terms of their attitudes and beliefs with respect to the value of work and work-related behavior. An important

consideration for industrial psychology is the relationship between these attitudes and beliefs and actual work behavior. While industrial psychologists interested in the work ethic have typically explored its relationship with other attitudinal variables such as job satisfaction (e.g., Aldag & Brief, 1975; Blood, 1969; Stone, 1975, 1976; Wanous, 1974), job involvement (e.g., Blau, 1987; Randall & Cote, 1991; Saal, 1978), and organizational commitment (e.g., Kidron, 1978; Morrow & McElroy, 1987), there have been relatively few studies (e.g., Khaleque, 1992; Orpen, 1986), focusing on the relationship of work ethic with actual job performance. A possible reason for this is the lack of distinction between task and contextual aspects of job performance.

Recently several models of job performance have been proposed which attempt to describe a set of underlying dimensions that are representative of performance in all jobs (Borman & Motowidlo, 1993; Campbell, 1990; Campbell, McCloy, Oppler, & Sager, 1993). For example, Campbell (1990) argues that all jobs are made up of eight factors, including: job-specific task proficiency, non-job-specific task proficiency, written and oral communication, demonstrating effort, maintaining personal discipline, facilitating team and peer performance, supervision and leadership, and management and administration. Campbell's formulation distinguishes between behaviors that contribute to organizational effectiveness through their focus on task proficiency and those behaviors that help the organization in other ways (Motowidlo & Van Scotter, 1994). Task proficiency behaviors are formally prescribed by the organization whereas other behaviors, though not formally a part of the job, are still very valuable for organizational effectiveness (Borman & Motowidlo, 1993).

Borman and Motowidlo (1993) place performance behaviors not prescribed by the organization under the rubric of contextual activities. Examples include:

(1) Volunteering to carry out task activities that are not formally a part of the job.
(2) Persisting with extra enthusiasm or effort when necessary to complete own task activities successfully.
(3) Helping and cooperating with others.
(4) Following organizational rules and procedures even when personally inconvenient.
(5) Endorsing, supporting, and defending organizational objectives. (p. 73)

Using a sample comprising Air Force mechanics, Motowidlo and Van Scotter (1994) demonstrated that supervisors consider task performance and contextual performance separately when providing performance ratings. It is the contextual component of job performance in which work ethic may offer substantial predictive utility.

Specifically, it may be possible to predict with a measure of work ethic the extent to which an individual would engage in contextual performance of value to the unit. Further, the work ethic may demonstrate a relationship with technical school training success, job performance, and tenure in the Air Force.

Measurement of Work Ethic

Of paramount concern for research focusing on the understanding of the work ethic construct as well as the relationship between work ethic and work behavior is the ability to accurately measure the construct. There are at least seven work ethic measures in existence which purport to provide reliable and valid measures of this construct. However, there are a number of problems with these measures. First and foremost, they focus on the measurement of a single construct by providing a global "work ethic" score. This is a considerable shortcoming as, since its inception, Weber believed the work ethic to be a multidimensional construct; a position that has subsequently been supported by numerous researchers (Bouma, 1973; Cherrington, 1980; Furnham, 1984; Oates, 1971).

From a psychometric as well as a conceptual perspective, the lack of focus on the multidimensional nature of the work ethic is troubling. The use of a single overall score could potentially cause the loss of information with regards to the different components of work ethic as well as their relationships with other constructs (Carver, 1989; McHoskey, 1994). Further, the use of a single score in studies using different instruments to measure the work ethic may at least partially explain the equivocal results often found in the literature (Furnham, 1984). That is, one cannot be sure if the conflicting results are due to a lack of robustness in the studies, the scales measuring different components of the work ethic, or deficiencies in terms of construct relevance and psychometric properties (Furnham, 1990b).

A second concern is that the various measures appear to tap different components of the work ethic and not the construct in its entirety. This has often led to poor intercorrelations among measures. For example, Furnham (1990b) administered seven measures of the work ethic to 1,021 participants and found that the correlations between the various measures ranged from 0.19 - 0.66 with a mean r of 0.36. One would expect the values to be much higher if the scales were indeed measuring the same thing.

Finally, another potential problem with existing work ethic measures is that these measures are relatively dated. The mean time since publication for the previous measures is 23 years. The age of the measures poses the

problem of many dated items. For example, some of the items contain sex-biased language such as: "Hard work makes a man a better person", "The man who can approach an unpleasant task with enthusiasm is the man who gets ahead", and "To be superior a man must stand alone".

Factor analytic investigations of the various measures have found the existence of several identifiable factors (Furnham, 1990b; Heaven, 1989; Tang, 1993; Mirels & Garrett, 1971; McHoskey, 1994). For example, McHoskey (1994) factor analyzed Mirels and Garrett's Protestant Ethic scale. His analysis yielded a 4-factor solution which he labeled, "success", "asceticism", "hard-work", and "anti-leisure". However, McHoskey was quick to point out that though this scale was multidimensional, other important aspects of the PWE were absent. Specifically, it in no way measured an individual's attitudes toward morality, self-reliance, or delay of gratification. This lack of comprehensiveness in measuring the work ethic has been levied against other scales as well and limits their utility (Furnham, 1984, 1990a, b; McHoskey, 1994).

In an effort to ameliorate the shortcomings in previous attempts to measure the work ethic, Woehr and Miller (1997) and Miller and Woehr (1997) developed the Multidimensional Work Ethic Profile (MWEP). The goal in the development of such a measure was to build on and extend previous measures in an attempt to capture the multidimensionality of the construct. The MWEP is a 65-item measure assessing 7 dimensions related to the work ethic construct. These dimensions are: *"Delay of Gratification"*, *"Hard Work"*, *"Morality/Ethics"*, *"Self-Reliance"*, *"Leisure"*, *"Wasted Time"*, and *"Centrality of Work"*. Complete definitions of these dimensions are provided in table 1.

Originally developed based on a sample of university students, the MWEP has demonstrated good psychometric characteristics including reliability and validity. Specifically, Miller and Woehr (1997) report 3 - 4 week test-retest reliabilities of 0.83 - 0.95 and internal consistency coefficient alphas of 0.78 - 0.89 for the dimensions of work ethic. With regards to construct-related validity the MWEP demonstrated discriminant relationships with personality, cognitive ability, and manifest needs. Lastly, the criterion-related validity of the MWEP was evaluated by relating it to academic effort indices pertinent to the university student sample. The MWEP was shown to be significantly related to hours studying per week (0.21), hours watching TV per week (0.36), hours in extracurricular activities per week (0.26), and classes missed (0.30).

Table 1.

Dimension definitions for the 7 work ethic dimensions assessed by the MWEP.

Dimension:	**Definition:**
Centrality of Work	Belief in the virtues of hard work.
Delay of Gratification	Striving for independence in one's daily work.
Hard Work	Pro-leisure attitudes and beliefs in activities that serve a rejuvenating function.
Leisure	Belief in work for work's sake and the importance of work.
Morality/Ethics	Believing in a just and moral existence.
Self-Reliance	Orientation toward the future; the postponement of rewards.
Wasted Time	Attitudes and beliefs reflecting active and productive use of time.

Present Study

Given the previous evaluations of the MWEP and the potential for use as a screening measure among Air Force enlisted personnel, the objective of this study was to empirically determine the extent to which the psychometric properties of the MWEP that have been found with a university student sample would generalize to Air Force enlisted personnel. Measurement stability across the samples would allow for greater confidence with regards to measurement equivalence and provide an initial indication of the viability of the MWEP for use in the Air Force.

As noted, the primary objective of the present study was to compare the psychometric characteristics of the MWEP with Air Force personnel relative to the original student development sample. This comparison focused on: (1) the mean score levels on each dimension, (2) score variability for each dimension, (3) the reliability for each dimension, and (4) the overall pattern of correlations among dimensions. If the MWEP functions similarly across the two samples no differences in dimension variability, dimension reliability, or the overall pattern of correlations among dimensions should be found. However, differences in mean levels on each dimension are likely given the actual differences across the two samples. That is, the student sample represents 18 to 22-year-old college students. Alternately, the Air Force sample represents an 18 to 22-year-old non-college bound sample. It is likely that actual

differences in work ethic attitudes and beliefs exist across the two groups. Such differences would be reflected in mean dimension score differences.

Method

University Participants.

The university student sample comprised 598 participants (52% female and 48% male). Subject participation was voluntary and subjects received partial course credit for taking part in the study. Mean age of the participants was 19.2 and ranged from 17 to 27.

Air Force Participants.

Participants in the present study were 268 Air Force enlisted personnel that participated in the study during Basic Military Training (BMT). The participants were 95% male and 5% female. Further, 71% were White, 15% Black, 8% Hispanic, 4% Asian, and 2% Other. Mean age of the participants was 19.95 and ranged from 18 to 35.

Multidimensional Work Ethic Profile (MWEP) Measure.

The MWEP was originally developed as a 65 item paper-and-pencil measure. The measure requires responses to items on a 5 point Likert-type scale ranging from 1(strongly disagree) to 5 (strongly agree). In order to facilitate data collection in the present study the MWEP was included as part of a computer-administered battery of questionnaires. Thus a computer administered version of the MWEP was developed. Although computer-administered this version was highly similar to the paper-and-pencil version. Both items and response options were displayed in the same manner in both forms. Participants were asked to respond to each of the items via the numbers on the computer keyboard.

Procedure.

The MWEP was administered as part of an extensive battery of computer-administered questionnaires completed in a single 4 hour session during the first week of BMT. Subjects were seated at individual computer terminals and given the measures. Administration of the measures was counterbalanced across experimental sessions.

Results

Comparison of the MWEP in the two samples focused on: (1) the mean score levels on each dimension, (2) score variability for each dimension, (3) the reliability for each dimension, and (4) the overall pattern of correlations among dimensions. Mean scores for each of the 7 work ethic dimensions for both the Air Force and student samples are presented in Table 2.

Table 2.

Means and standard deviations for the 7 work ethic dimensions for both the Student and Air Force Samples.

	Student Sample N = 598		**Air Force Sample N = 268**		
Dimension:	**Mean**	**SD**	**Mean**	**SD**	**t**
Centrality of Work	24.34	6.04	20.79	6.03	8.01*
Delay of Gratification	16.95	4.52	14.08	4.18	8.85*
Hard Work	22.09	5.86	16.81	5.45	12.52*
Leisure	28.63	5.86	31.50	5.81	-6.68*
Morality/Ethics	16.08	4.45	13.90	3.21	8.15*
Self-Reliance	26.11	6.88	24.84	6.79	2.51
Wasted Time	19.96	4.71	16.10	4.34	11.41*

* $p < .01$.

Tests for differences between the mean scores for each dimension are also presented in Table 2. These results indicate significant mean differences for all dimensions except the self-reliance dimension. Further, means are higher for the student sample than for the Air Force sample for all dimensions except the leisure dimension. Thus, the student sample had significantly higher mean scores for the *"Centrality of Work"*, *"Delay of Gratification"*, *"Hard Work"*, *"Morality/Ethics"*, and *"Wasted Time"* dimensions. The Air Force sample had a significantly higher mean score on the leisure dimension and no significant difference was for the self-reliance dimension.

Table 3 provides the results of a comparison of the variance of each dimension across samples. These results indicate no significant differences for any of the dimensions across samples except *"Morality/Ethics"*. For the morality/ethics dimension there is significantly less variability in scores for the Air Force sample than for the student sample.

Table 3.

Test for equality of variances across student and Air Force Samples.

			Levine's Test for Equality of Variances	
Dimension	**Student Sample Variance**	**Air Force Sample Variance**	**F**	**Significance**
Centrality of Work	36.48	36.36	.088	.767
Delay of Gratification	20.43	17.47	2.60	.107
Hard Work	34.33	29.70	1.173	.279
Leisure	34.33	33.76	.187	.665
Morality/ethics	19.80	10.30	26.301	.000
Self-Reliance	47.33	46.10	.136	.713
Wasted Time	22.18	18.84	1.586	.208

Dimension reliabilities (coefficient α) for both samples are presented in table 4. Examination of these results indicate no differences in dimension reliabilities across samples except for the *"Morality/Ethics"* dimension. Specifically, all dimension reliabilities are within .03 of each other across samples except for the *"Morality/Ethics"* dimension for which the reliability is substantially lower in the Air Force sample.

Finally, the dimension intercorrelations for both the Air Force and student samples are presented in table 5. In order to assess the extent to which the dimension intercorrelations differed across samples, we used LISREL 8.14 (Joreskog & Sorbom, 1993) to provide an overall test of the equivalence of the 2 correlation matrices. Specifically,

we tested a model in which correlations among the 7 work ethic dimensions were set equal to the student sample based correlations and the correlations for the Air Force sample were constrained to be equal to those from the student sample. Using this approach, the overall model fit indices derived from the LISREL analyses provide an indication of the overall equality of the correlations across samples. Results of this analysis are provided in Table 6 and indicate that the two sets of correlations are generally equivalent.

Table 4.

Test for equality of reliabilities across student and Air Force Samples.

Dimension	Student Sample Reliability	Air Force Sample Reliability
Centrality of Work	.84	.85
Delay of Gratification	.79	.78
Hard Work	.85	.87
Leisure	.87	.87
Morality/ethics	.78	.55
Self-Reliance	.89	.86
Wasted Time	.79	.77

Discussion

The present study presents an examination of the psychometric properties of the Multidimensional Work Ethic Profile (MWEP) developed by Michael Miller and David Woehr (Woehr & Miller, 1997, Miller and Woehr, 1997). The MWEP is a 65 item measure of work ethic based on previous research and literature focusing on work ethic and job performance. An important characteristic of the MWEP is that it assess 7 conceptually and empirically distinct facets of the work ethic construct. Originally developed based on a sample of university students, the MWEP has demonstrated good psychometric characteristics including reliability and convergent and

discriminate validity. Further, the MWEP has been suggested as a potentially valuable screening tool with Air

Table 5.

Work ethic dimension intercorrelations for the student and Air Force samples.

	Student Sample						
Dimensions:	1	2	3	4	5	6	7
1. Centrality of Work	1.0						
2. Delay of Gratification	.38	1.0					
3. Hard Work	.33	.33	1.0				
4. Leisure	-.47	-.12	-.08	1.0			
5. Morality/Ethics	.17	.25	.22	.08	1.0		
6. Self-Reliance	.20	.21	.38	.10	.13	1.0	
7. Wasted Time	.56	.40	.38	-.28	.21	.32	1.0
	Air Force Sample						
1. Centrality of Work	1.0						
2. Delay of Gratification	.52	1.0					
3. Hard Work	.48	.56	1.0				
4. Leisure	-.44	-.26	-.23	1.0			
5. Morality/Ethics	.34	.44	.48	-.16	1.0		
6. Self-Reliance	.20	.11	.23	-.02	.18	1.0	
7. Wasted Time	.62	.55	.59	-.34	.46	.26	1.0

Student Sample $N = 598$. All correlations are significant ($p < .01$).

Air Force Sample $N = 268$. All correlations greater than .17 are significant ($p < .01$).

Table 6.

Goodness of fit indices for the test of intercorrelation equivalence.

χ^2	df	χ^2/df	RMSEA	GFI	NFI	CFI	RFI
72.30	42	1.72	.04	.92	.95	.98	.95

Force enlisted personnel. The purpose of the present study was to provide a preliminary evaluation of the measure among Air Force enlisted personnel. Results indicate that the measure does in fact demonstrate highly similar psychometric characteristics among Air Force enlisted personnel as with the original developmental sample. The MWEP provides reliable and valid measures of multiple dimensions underlying the work ethic construct. These results indicate that the MWEP may be a useful screening tool for Air Force Personnel.

Specifically, results of the present study found no differences across samples for the dimension variances, reliabilities, and intercorrelations across dimensions. One exception to these findings was for the *"Morality/Ethics"* dimension. For this dimension the results indicated significantly less variance as well as substantially lower reliability with the Air Force sample relative to the student sample. One possible explanation for this finding may lie in differences in the work settings of the two samples. That is, the student sample was assessed in a non-job setting while the Air Force sample was assessed in an actual job setting. It is likely that the items comprising the *"Morality/Ethics"* dimension are fairly transparent and actual job incumbents may not respond as truthfully as non incumbents. This would explain the restricted variance found in the Air Force sample. This reduced variance would in turn result in a lower reliability estimate. Counter to this explanation, however, was our finding that the mean response for the *"Morality/Ethics"* dimension was actually significantly lower in the Air Force sample relative to the student sample. If the items were relatively transparent and the incumbent sample was simply responding in a more socially desirable manner then one would expect a higher mean score. It is difficult at this point to determine the exact reasons for the differences found across samples for this dimensions. The lack of differences across the other, more work-related, dimensions however is encouraging.

The results of the present study do indicate significant mean score differences for 6 of the 7 dimensions across samples. These differences are not unexpected and do not call into question the measurement equivalence of the MWEP in either sample. Rather these differences are to a certain extent consistent with expected differences between the two samples. The student sample represents young adults attending college. Alternately, the Air Force sample represents young adults not attending college but directly entering the work force. Thus differences in work ethic scores most likely reflect actual differences between samples.

Conclusion

The prediction of job performance is one of the benchmarks of industrial psychology. Though the field has relied primarily on cognitive ability measures to predict performance, it has also pursued the use of alternative predictors (Arvey & Sackett, 1993). One of the most prevalent alternative predictors has been personality variables (Adler, 1996; Barrick & Mount, 1991; Goffin, Rothstein, & Johnston, 1996; Hogan, Hogan, & Roberts, 1996; Horman & Maschke, 1996; Tett, Jackson, & Rothstein, 1991). Though measures of personality have not resulted in adverse impact, many researchers have found a low relationship with actual criterion measures of job performance (Ones, Mount, Barrick, & Hunter, 1994). Another potential problem is that personality variables may not function in a linear fashion. Attitudinal variables such as work ethic may bridge the gap between cognitive ability and

personality variables.

The present study demonstrates that one such attitudinal measure, the MWEP, a multidimensional measure of work demonstrates good psychometric characteristics in two diverse samples. This suggests that the MWEP is a potentially valuable pragmatic measure for either sample. Certainly, the next step is to examine the predictive utility of the MWEP in an Air Force context. A proposed avenue of research for the future would be an examination of the relationship of the work ethic to technical school training success, job performance, and tenure in the Air Force. This could be achieved through the administration of the MWEP to enlisted personnel while in BMT and following up on their respective progress in the Air Force. The criteria in this example might be technical school final grades, performance evaluations while at the duty station, and fulfilment of enlistment tour requirements.

References

Aldag, R. J. & Brief, A. P. (1975). Some correlates of work values. Journal of Applied Psychology, 60, 757-760.

Blau, G. J. (1987). Using a person-environment fit model to predict job involvement and organizational commitment. Journal of Vocational Behavior, 30, 240-257.

Blood, M. R. (1969). Work values and job satisfaction. Journal of Applied Psychology, 53, 456-459.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations (pp. 71-98). San Francisco: Jossey-Bass.

Bouma, G. D. (1973). Beyond Lenski: A critical review of recent "Protestant Ethic" research. Journal for the Scientific Study of Religion, 12, 141-155.

Byrne, E. F. (1990). Work, inc.: A philosophical inquiry. Philadelphia: Temple University Press.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (Vol 1, 2nd ed., pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations (pp. 35-70). San Francisco: Jossey-Bass.

Carver, C. S. (1989). How should multifaceted personality constructs be tested? Journal of Personality and Social Psychology, 56, 577-585.

Charlton, W., Mallinson, T., & Oakeshott, R. (1986). The Christian response to industrial capitalism. London: Sheed & Ward.

Cherrington, D. J. (1980). The work ethic: Working values and values that work. New York: AMACOM.

Dubin, R. (1963). Industrial worker's worlds: A study of the 'central life interests' of industrial workers. In E. O. Smigel (Ed.), Work and Leisure (pp. 53-72). New Haven: College and University Press.

Fine, R. (1983). The Protestant ethic and the analytic ideal. Political Psychology, 4, 245-264.

Furnham, A. (1984). The Protestant work ethic: A review of the psychological literature. European Journal of Social Psychology, 14, 87-104.

Furnham, A. (1990a). The Protestant work ethic: The psychology of work-related beliefs and behaviours. London: Routledge.

Furnham, A. (1990b). A content, correlational, and factor analytic study of seven questionnaire measures of the Protestant work ethic. Human Relations, 43, 383-399.

Gilbert, J. B. (1977). Work without salvation: America's intellectuals and industrial alienation, 1880-1910. Baltimore: Johns Hopkins University Press.

Giorgi, L., & Marsh, C. (1990). The Protestant work ethic as a cultural phenomenon. European Journal of Social Psychology, 20, 499-517.

Green, A. W. (1968). Sociology: An analysis of life in modern society. New York: McGraw-Hill.

Ho, R., & Lloyd, J. I. (1984). Development of an Australian work ethic scale. Australian Psychologist, 19, 321-332.

Heaven, P. C. L. (1989). Structure and personality correlates of the Protestant work ethic among women. Personality and Individual Differences, 10, 101-104.

Joreskog, K. G., Sorbom, D. (1993). LISREL-8: Structural equation modeling with the SIMPLIS command language. Chicago, IL: Scientific Software International Inc.

Khaleque, A. (1992). Work values, attitudes and performance of industrial workers in Bangladesh. Social Indicators Research, 27, 187-195.

Kidron, A. (1978). Work values and organizational commitment. Academy of Management Journal, 21, 239-247.

Lehmann, H. (1993). The rise of capitalism: Weber versus Sombart. In H. Lehmann & G. Roth (Eds.), Weber's Protestant ethic: Origins, Evidence, Contexts (pp. 195-208). Washington, D. C.: Cambridge University Press.

Lessnoff, M. H. (1994). The spirit of capitalism and the Protestant ethic: An Enquiry into the Weber thesis. Brookfield, VT: E. Elgar.

Maccoby, M. (1983). The managerial work ethic in America. In J. Barbash, R. J. Lapman, S. A. Levitan, & G. Tyler (Eds.), The work ethic--A critical analysis (pp. 183-196). Madison, WI: Industrial Relations Research Association.

McHoskey, J. W. (1994). Factor structure of the Protestant work ethic scale. Personality and Individual Differences, 17, 49-52.

Mirels, H. L., & Garrett, J. B. (1971). The Protestant ethic as a personality variable. Journal of Consulting and Clinical Psychology, 36, 40-44.

Morrow, P. C., & McElroy, J. C. (1987). Work commitment and job satisfaction over three career stages. Journal of Vocational Behavior, 30, 330-346.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. Journal of Applied Psychology, 79, 475-480.

Nelson, B. (1973). WeberÕs Protestant ethic: Its origins, wanderings, and foreseeable future. In C. Glock & P. Hammond (Eds.), Beyond the classics (pp. 71-130). New York: Harper & Row.

Nord, W. R., Brief, A. P., Atieh, J. M., & Doherty, E. M. (1988). Work values and the conduct of organizational behavior. Research in Organizational Behavior, 10, 1-42.

Poggi, G. (1983). Calvinism and the capitalist spirit: Max Weber's Protestant ethic. Amherst: University of Massachusetts Press.

Miller, M. J., & Woehr, D. J. (1997). Work ethic: The development and evaluation of a multidimensional measure. Manuscript submitted for publication.

Oates, W. (1971). Confessions of a workaholic: The facts about work addiction. New York: World Publishing Company.

Pascarella, P. (1984). The new achievers: Creating a modern work ethic. New York: Free Press.

Randall, D. M. & Cote, J. A. (1991). Interrelationships of work commitment constructs. Work and Occupations, 18, 194-211.

Ray, J. J. (1982). The Protestant ethic in Australia. Journal of Social Psychology, 116, 127-138.

Saal, F. E. (1978). Job involvement: A multivariate approach. Journal of Applied Psychology, 63, 53-61.

Stone, E. F. (1975). Job scope, job satisfaction, and the Protestant ethic: A study of enlisted men in the U.S. Navy. Journal of Vocational Behavior, 7, 215-234.

Stone, E. F. (1976). The moderating effect of work-related values on the job scope-job satisfaction relationship. Organizational Behavior and Human Performance, 15, 147-167.

Tang, T. L. -P. (1993). A factor analytic study of the Protestant work ethic. The Journal of Social Psychology, 133, 109-111.

Wanous, J. P. (1974). Individual differences and reactions to job characteristics. Journal of Applied Psychology, 59, 616-622.

Weber, M. (1958). The Protestant Ethic and the spirit of capitalism (T. Parsons, Trans.). New York: Charles Scribner's Sons.

Woehr, D. J. & Miller, M. J. (1997, April). The meaning and measurement of work ethic. Paper presented at the annual Southeastern Industrial/Organizational Psychology Meeting, Atlanta, GA.

Wollack, S., Goodale, J. G., Wijting, J. P., & Smith, P.C. (1971). Development of the Survey of Work Values. Journal of Applied Psychology, 55, 331-338.

THE EFFECT OF SIZE DISPARITY ON PERCEPTION OF SURFACE SLANT IN STEREOSCOPIC MOVING IMAGES

Miguel A. Moreno
Graduate Assistant
Department of Psychology

Arizona State University
Tempe, AZ

Final Report for:
Summer Graduate Research Program
Armstrong Laboratories

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

Armstrong Laboratory

August 1997

THE EFFECT OF SIZE DISPARITY ON PERCEPTION OF SURFACE SLANT IN MOTION CONTEXTS

Miguel A. Moreno
Graduate Assistant
Department of Psychology
Arizona State University

Abstract

This study focuses on the effects of horizontal, vertical and overall-size disparities on perceived surface slant under both static and motion conditions. The effect of superimposed zero-disparity contrast stimuli was also measured. The response patterns from all test conditions show no distinction between moving versus static disparity images. The results corroborate the previous findings by Pierce & Howard (1996) and Kaneko & Howard (1996) that a superimposed zero-disparity contrasting image effects the perceived inclination of the disparity surface. New interactions between vertical size disparities and a superimposed zero-disparity horizontal line are presented.

THE EFFECT OF SIZE DISPARITY ON PERCEPTION OF SURFACE SLANT IN MOTION CONTEXTS

Miguel A. Moreno

Introduction

Ogle (1938) reported that a difference in the sizes of the images in the two eyes produced a perceived surface slanted in depth about a vertical axis. He showed that if an image is horizontally magnified to one eye the perceived surface appears to be rotated farther away on the side with the magnified image. The theoretical relation between slant angle and horizontal-size ratio is:

$$\tan \theta = 2D\ (R-1)/I(R+1)$$

where θ is the angle of the slant, R is the horizontal-size ratio of the two images, D is the observation distance , and I is the interpupilary distance.

Ogle also showed that if an image is similarly magnified in the vertical dimension, the perceived surface appears rotated nearer on the side with the magnified image. He called this phenomenon the induced effect. Koenderink and van Doorn (1976) proposed that this phenomenon depends on the deformation component of disparity rather than on horizontal disparity alone. The deformation component is defined as the local difference between the horizontal-size disparity and the vertical size disparity.

Mayhew and Longuet-Higgins (1982) realized that vertical disparities in a frontal surface increase with lateral eccentricity and decrease as a function of the absolute distance of the surface. This lead them to conclude that the effects of vertical disparity could be used to judge absolute distances. Rogers and Bradshaw (1993, 1995) provided empirical evidence using larger surfaces which show that the ratio of horizontal to vertical disparities can be used to make judgments whether a surface lies in the frontal plane. This would suggest that disparity information is processed more globally rather than locally.

Kaneko and Howard (1996) showed that viewing horizontal and overall-size disparity with a zero-disparity surround enhances the perceived slant of the central

display. For vertical-size disparities, the presence of a zero-disparity surround reduced the perceived slant of the central display. Kaneko and Howard also found that a random-dot stereograms with horizontal disparities superimposed with a zero-disparity pattern of random dots appear as two distinct surfaces whereas a random-dot stereograms with vertical disparities superimposed with a zero-disparity pattern of random dots appeared as one plane. This suggest the strategies for processing horizontal disparities are distinct from processing strategies associated with vertical disparities. Kaneko & Howard concluded that horizontal disparities may are processed locally whereas vertical disparities are integrated globally. This hypothesis was supported by Pierce and Howard (1996) who concluded that perceived slant of disparity surfaces and stimuli containing zero-disparity is assessed at each location in terms of the difference between the local horizontal-size disparity and the vertical-size disparity averaged over a large region.

Hadani and Vardi (1987) and Vardi and Hadani (1989) reported an inhibitory effect of motion on the depth perception of a pattern of random dots across a stationary stereograting. The authors attribute this impairment to a critical alternation rate, above which the two depth planes were averaged together creating the perception of one depth plane. They argued that the cause of the impairment is a temporal integration process which averages the alternating disparity values of the moving dots.

The present experiment is designed to investigate an inhibitory effects of motion on horizontal, vertical and overall-size disparity patterns. We are particularly interested how previously reported interactions between size disparity patterns and zero-disparity stimuli are affected by a motion context.

Methods

Participants

Three males with normal non-corrected vision participated in this study. All participants were paid.

Stimuli

Images were displayed on two computer monitors (Radius model TX-D2151RD running at 60 Hz) mounted in a mirror haploscope consisting of two 40% reflective mirrors (60% transmissive). The two mirrors were positioned at a 45° angle with respect to the observers line of sight. The distance from the vertex of the 90° angle formed by the two glass planes and the center of the monitors was 33.5 cm. The mirrors were mounted at a height to allow for line of sight to be centered with respects to the center of the monitors.

Responses to the stimuli were made using a response disk, 17.5 cm in diameter and 6 mm thick. The paddle could be to rotate about a vertical axis. The response disk was mounted directly in front of the observer at the same distance as the perceived surface (33.5 cm). The disk was centered with respects to the binocularly fused image.

Each stimuli pattern was a computer-generated pattern of elements consisting of open squares, crosses and lines. The squares and crosses subtended 0.85° of visual angle in both the horizontal and vertical dimensions. The lines also subtended a 0.85° visual angle in either a vertical or horizontal dimension. The lines for the elements were 0.083° in width and red which has a lower rate of phosphor decay. The elements were randomly distributed in a 10 by 14 matrix. The disparity pattern was presented either in isolation or with one of two superimposed zero-disparity images. The superimposed zero-disparity stimuli consisted of either a horizontal line subtending 15° of visual angle or a randomly distributed pattern of elements. The zero-disparity line was centered with respects to the disparity image. The zero-disparity pattern contained the same set of texture elements as the disparity pattern which were randomly distributed about an 11 by 14 matrix. A subtle change in light intensity of the elements in the zero-disparity pattern every 5 second assisted the observer to distinguish them from the elements in the disparity

pattern. Both the rows and columns of the zero-disparity pattern alternated with the rows and columns of the disparity pattern.

The same pattern of elements was displayed on the two computer monitors of the mirror haploscope. The fused images appeared in the median plane of the observer at a distance of approximately 33.5 cm. The fused display subtended a 53° horizontally and 43° vertically. All other perspective cues were eliminated and the surroundings were black.

The size disparity was introduced into the image by increasing the size of the image shown to either the right or left eye. Three type of size disparity were shown: horizontal, vertical and overall, as shown in figure 1. In both the horizontal and vertical size disparity conditions the spacing between the elements as well as the size of each element was increase along one dimension. For overall-size disparity, the spacing and size of each element was increased along both dimensions. Four levels of horizontal, vertical and overall-size disparity was used: 2%, 4%, 6% and 8% with a control level of 0%. Each of the four levels of size disparity was presented to both the right and left eye (this manipulation will be referred to as direction of disparity, positive for right eye and negative for left eye).

Two type of motions were used, downward and across (right to left) at a rate of 4° per second. In the trials with a zero-disparity pattern, both sets of elements moved in the same direction and at the same rate.

Insert Figure 1 about here

A separate calibration procedure used a 33.5 X 26.8 cm surface with a randomly distributed pattern of elements. A computer generated pattern of elements was printed and mounted on a stiff board. The surface contained the same set of texture elements and density of elements as the disparity patterns. The surface of texture elements was place

at the same distance from the observer as the perceived test image. The slant of the surface could be adjusted to any angle about a vertical axis.

Procedure

Participants were first prescreened for both visual acuity and stereo deficits. A red-green Snellen chart was used to test visual acuity and a Randot™ stereo test was used to detect stereo deficiencies. During the prescreening, each observers interpupilary distance (IPD) was measured using an Essilor™ Instruments digital corneal reflection pupilometer.

A block of trials consisted of 3 disparity types X 3 motion types X 4 disparity levels X 2 disparity directions X 3 levels of superimposed zero-disparity stimuli for a total of 216 test trials. A block to trials was divided into 9 trial sets by motion type and superimposed zero-disparity stimuli. Each set of 24 test trials was randomly presented along with one control trial with a disparity level of 0. Each participant responded to 3 blocks of trials. The sequence of the presentation of each trial set with in a block was determined based on a Latin square design.

The subject were tested over a nine day period. At the beginning of each participants were given the same set of 12 practice trials to allow for sufficient adaptation to the context. A total of three trial sets were presented on each day with a half-hour of rest between each set to relieve any accommodative demands.

During a trial presentation, the participants were first ask to report the number of columns and rows that appeared difficult to fuse. The observers were then asked to make a general description of the image and estimate the amount of slant perceived. Participants were given verbal instructions to observe the central 20° region of the display. The observer then adjust the unseen response disk to match the slant of the disparity pattern. Observers then adjusted the disk to match the perceived slant of the zero-disparity stimulus if one was presented.

In a separate control procedure, participants set the unseen response disk to match the slant of a surface of texture elements. The surface was slanted randomly at 5° intervals between ± 60° with a full range of binocular and monocular depth cues. For each angle, observers provided four consecutive responses with the paddle being reset to 0° between each responses. The results were fitted with a third-order polynomial function for each participant and used to calibrate the disk responses in the experiment.

Results

Perceived slant for horizontal-size disparity

Preliminary mean results for two of the three subjects for the horizontal-size disparity conditions are shown in figure 2. Each abscissa is the total relative horizontal-size disparity in the disparity pattern with the positive values signifying that the right eye's image was magnified. Each ordinate is the mean of the transformed setting of the response disk for the three subjects. Values above and below the zero line indicate that the fused image was perceived as right side away and left side away respectively.

Insert Figure 2 about here

For static, across, and down motion context the theoretically predicted function accounts for the pattern of data remarkably well. The fact that in each of the figures all three lines are virtually identical, indicates that motion has no significant effect on the perceived inclination of slant for horizontal size -disparity images. The perception of the surface was predictably perceived as slanted farther way on the side with the magnified image. Therefore, if the magnification was in the right eye, the image looked slanted right side away about a vertical axis. The polynomial function-like shape of the combined lines indicates a strong main effect of the disparity size. The amount of perceived slant was dependent on the percent magnification in the image disparity, particularly at lower

disparity sizes. Due to the symmetry of the plotted lines above and below the 0° line there does not seem to be an effect of direction of disparity. Comparing at the average perceived inclination collapsed across the static, across and down lines between figures 2a, 2b and 2d, there is some evidence for a slight contrast effect due to the presence of a superimposed zero-disparity image. The perceived inclination of the horizontal-size disparity image was greater when a zero-disparity pattern was also presented then the disparity pattern was presented in isolation. The effect of the presence of a zero-disparity pattern was most clearly evident at larger disparity size values. The difference between the perceived inclination of the horizontal-size disparity pattern with a line does not seem to be significantly different from either the horizontal-size disparity pattern alone or the horizontal-size disparity pattern with a pattern.

Comparing the responses to the zero-disparity stimuli shown in figure 2c and 2e, results show a significant difference between responses to the line versus responses to the pattern. In neither figure does there appear to be any separation of the lines indicating minimal effects of motion. The responses to the zero-disparity line appear to be an inverse of the corresponding responses to the horizontal-size disparity pattern. This indicates that the observers reported the line to be slanted at an angle opposite to that of the perceived disparity plane. The perception of the line also shows a sensitivity to changes in disparity size. Most observers reported the line to appear to intersect the disparity plane. The superimposed zero-disparity pattern was consistently perceived as being in the frontal plane. Neither, the zero-disparity line or pattern show any effect of direction of the disparity surface.

Participants reported no problem in fusing the whole of the pattern with horizontal-size disparity. Observers also fused the zero-disparity line and all elements of the zero-disparity pattern when presented. In the superimposed stimuli conditions, the zero-disparity stimuli and the surface with horizontal-size disparity appeared to lie on distinct planes.

Perceived slant for Vertical-size disparity

Preliminary mean results for two of the three participants for the vertical-size disparity conditions are presented in figure 3. Similar to the mean results for the horizontal size-disparity pattern, the lack of separation between the lines indicates no effect of motion. Figure 3a shows that vertical disparity patterns viewed in isolation appeared to slant in a direction opposite to that of a surface with the same sign of horizontal-size disparity. This induced effect was a function of direction and magnitude of vertical-size disparity. This was also true of vertical-size disparity patterns view with a zero-disparity line. The interaction between the perceived inclination of the zero-disparity line at various levels of size-disparity in the disparity surface. A superimposed zero-disparity pattern reduced the perceived slant of the vertical disparity patterns to near zero.

Insert Figure 3 about here

The response to the zero-disparity stimuli does not seem to show any difference between the line and the pattern as seen in figures 3c and 3e. Both are perceived to lie on the frontal plane in all test conditions.

Two of the three subject reported having difficulty in fusing all rows of the pattern with greater than 4% magnification. The ability to fuse all the elements was reduced with the presence of a zero-disparity stimuli.

Perceived slant for overall-size disparity

Figure 4 shown the preliminary mean results for two of the overall-size disparity conditions. Comparing figure 4a, 4b and 4c shows that the perceived slant of the surface with overall-size disparity was enhanced by the superimposed zero-disparity pattern but

not by the zero disparity line. As seen in the both horizontal and vertical-size disparity conditions, there is minimal effects of motion in the data. The overall disparity appeared to slant in a direction similar to that of a surface with the same sign of horizontal-size disparity. Overall the magnitude of the response for overall-size disparity pattern was less than horizontal-size disparity patterns.

Insert Figure 4 about here

Comparing the responses to the zero-disparity stimuli shown in figure 4c and 4e, results show a significant difference between responses to the line versus responses to the pattern. In either figure does there appear to be any separation of the lines indicating an effect of motion. The response pattern of the zero-disparity line appears to be an inverse of the corresponding response to the overall-size disparity pattern. This indicates that the observers reported the line to be slanted at an angle opposite and equal to that of the perceived disparity plane. Line responses also show a sensitivty to changes in disparity size. Most observers reported that the line to appeared to intersect the disparity plane. The zero-disparity pattern was consistently perceived as being in the frontal plane. Neither, the zero-disparity line or pattern show any effect of direction of the disparity surface.

Observers could fuse the whole pattern with overall-size disparity when presented in isolation for lower levels of disparity size. For higher values of disparity size, participants reported difficulty in fusing 2-3 rows on the top and bottom of the image. The zero-disparity line had little effect on the ability to fuse the image. The addition of a zero-disparity pattern reduced the ability to fuse the image.

References

Hadani, I. & Vardi, N. (1987). Stereopsis impairment in apparently moving random dot patterns. Perception & Psychophysics, 42 (2), 158-165.

Kaneko, H. & Howard, I. P. (1996). Relative size disparities and the perception of surface slant. Vision Research,

Koenderink, J.J., & van Doorn, A.J. (1976). Geometry of binocular vision and a model of stereopsis. Biological Cybernetics, 21, 29-35.

Mayhew, J.E.W., & Longuet-Higgins, H.C. (1982). A computational model of binocular depth perception. Nature, 297, 376-378.

Ogle, K.N. (1938). Induced size effect I: A new phenomenon in binocular space-perception associated with the relative sizes of the images of the two eyes. AMA Archives of Ophthalmology, 20, 604-623.

Pierce, B. & Howard, I.P. (1996) Type of size disparity and the perception of surface slant. (in press)

Rogers, B.J., & Bradshaw, M.F. (1993). Vertical disparities differential perspective and binocular stereopsis. Nature, 361, 253-255.

Rogers, B.J., & Bradshaw, M.F. (1995). Disparity scaling and the perception of frontoparallel surfaces. Science, 221, 1409-1411.

Vardi, N., & Hadani, I. (1989). Stereo impairment during smooth pursuit eye tracking. Brain Behavior Evolution, 33, 99-103.

Captions

Figure 1.

Paired images used in the Size Disparity Conditions: (a) Horizontal-size, (b) Vertical-size, and (c) Overall-size disparity. The increase in the image size is shown in the light colored elements. Disparity pattern that those illustrated above were also presented with a superimposed zero-disparity line subtending 15° of visual angle and superimposed zero-disparity pattern of similar elements.

Figure 2.

Perceived slant of (a) horizontal size-disparity pattern by motion type, (b) horizontal size-disparity pattern with superimposed zero-disparity line by motion type, (c), zero-disparity line with horizontal-size disparity pattern by motion type (d) horizontal size-disparity pattern with superimposed zero-disparity pattern by motion type, and (f) zero-disparity pattern with horizontal-size disparity pattern by motion type. Mean results for 3 subjects. The solid line is the theoretical slant of the disparity pattern as a function of its horizontal disparity.

Figure 3.

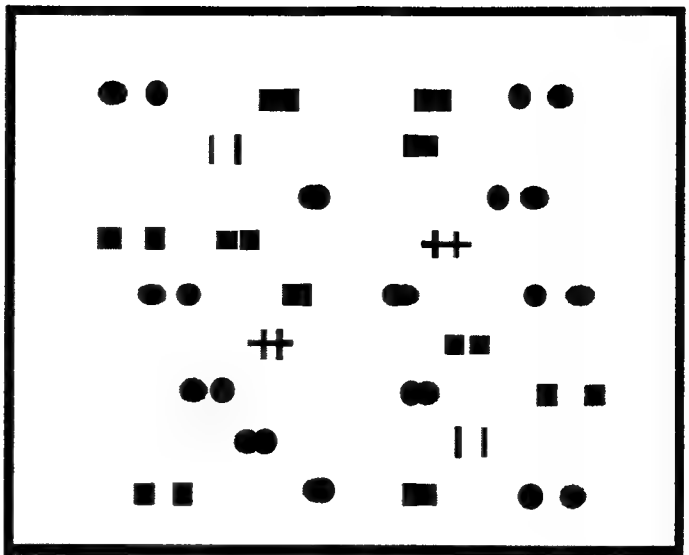
Perceived slant of (a) vertical size-disparity pattern by motion type, (b) vertical size-disparity pattern with superimposed zero-disparity line by motion type, (c), zero-disparity line with vertical-size disparity pattern by motion type (d) vertical size-disparity pattern with superimposed zero-disparity pattern by motion type, and (f) zero-disparity pattern with vertical-size disparity pattern by motion type. Mean results for 3 subjects.

Figure 4.

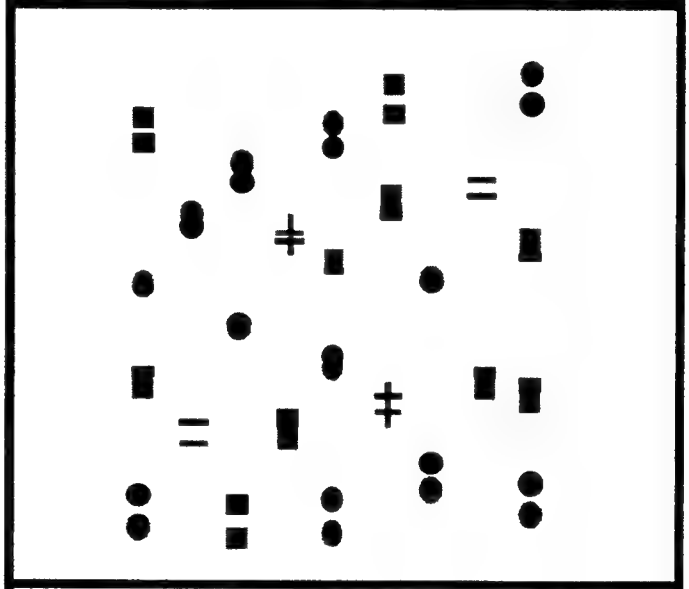
Perceived slant of (a) overall size-disparity pattern by motion type, (b) overall size-disparity pattern with superimposed zero-disparity line by motion type, (c), zero-disparity line with overall-size disparity pattern by motion type (d) overall size-disparity pattern with superimposed zero-disparity pattern by motion type, and (f)

zero-disparity pattern with overall-size disparity pattern by motion type. Mean results for 3 subjects

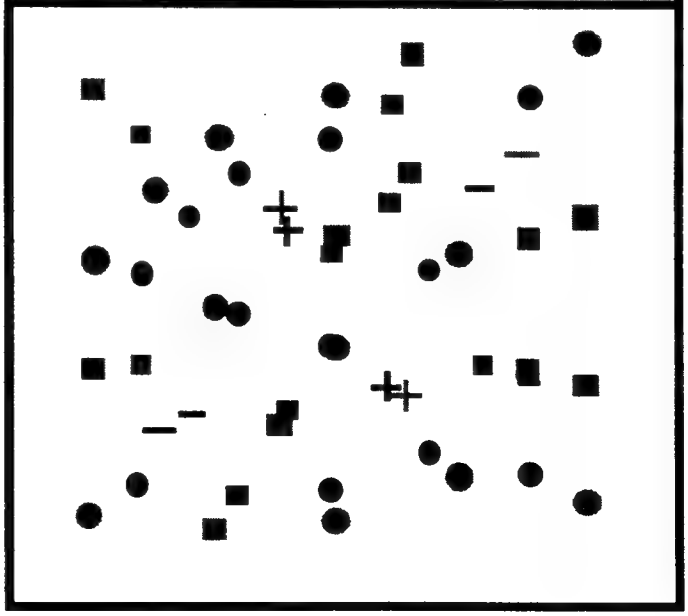
A

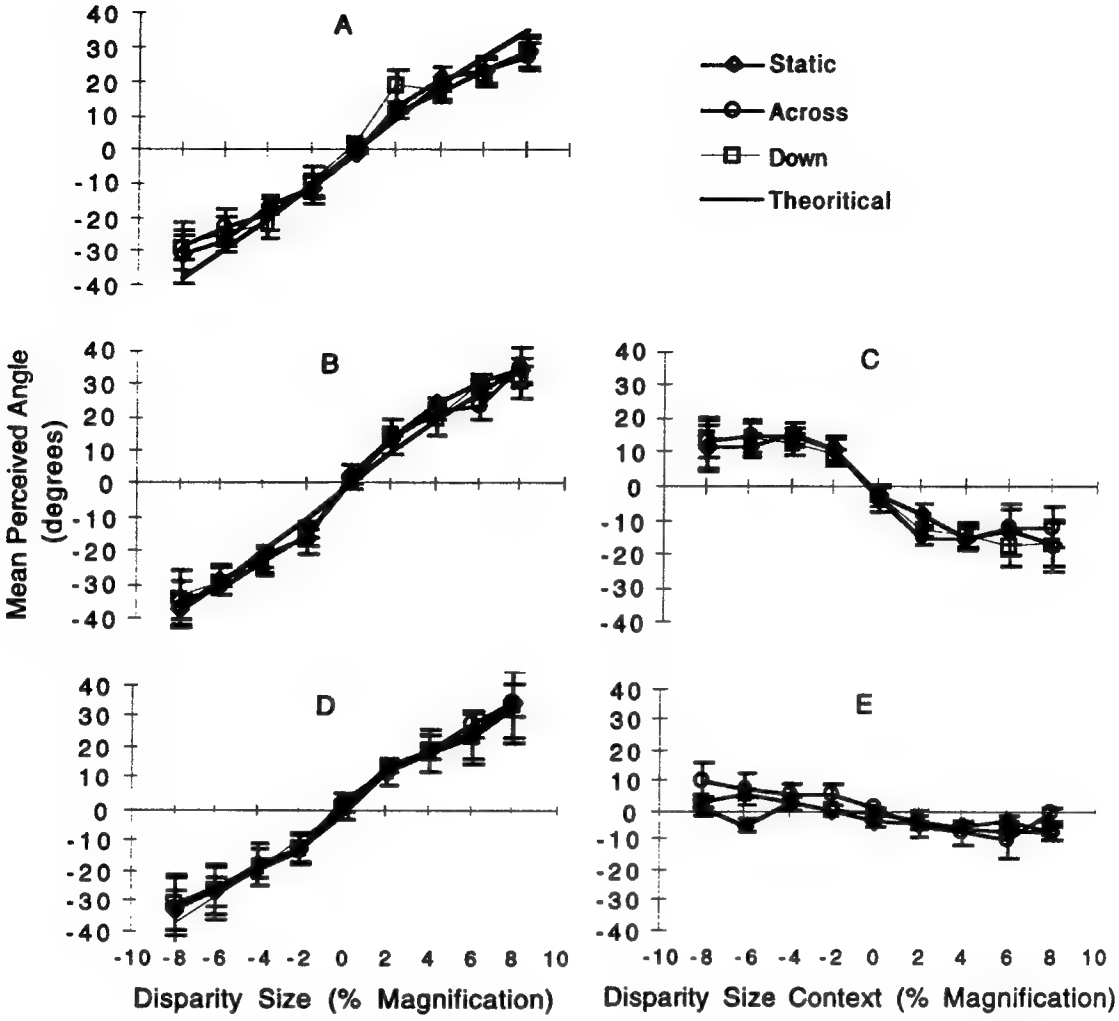


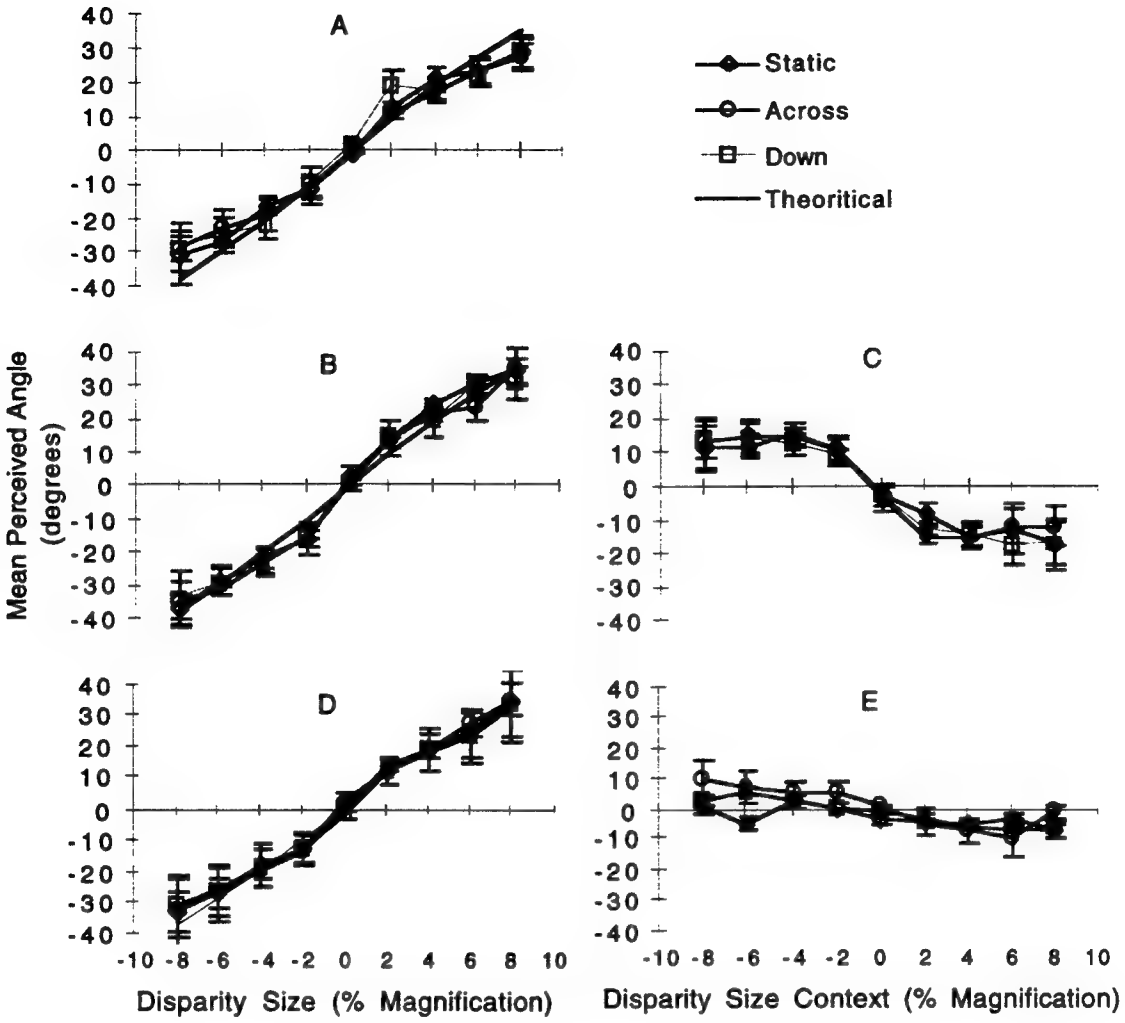
B

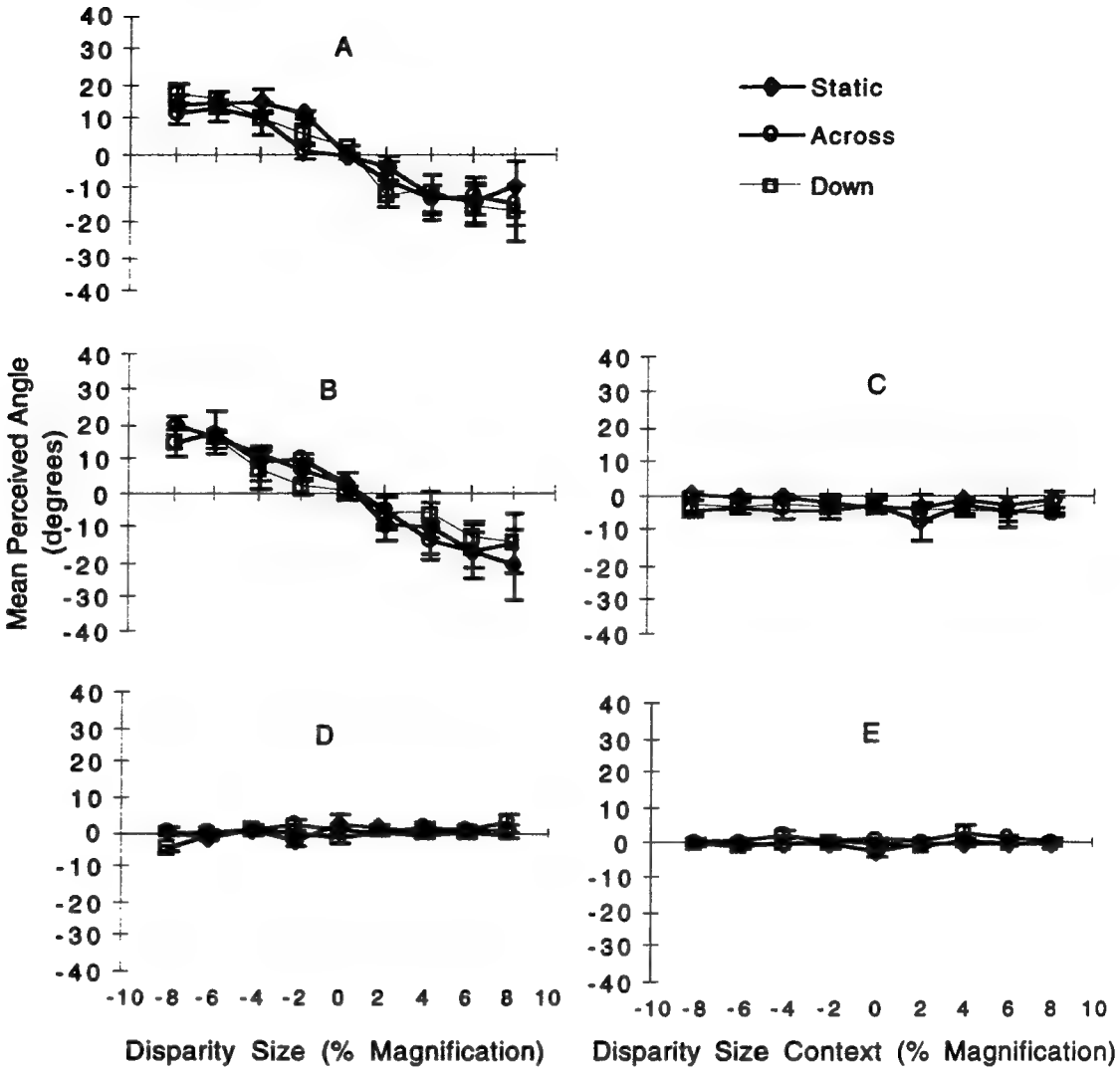


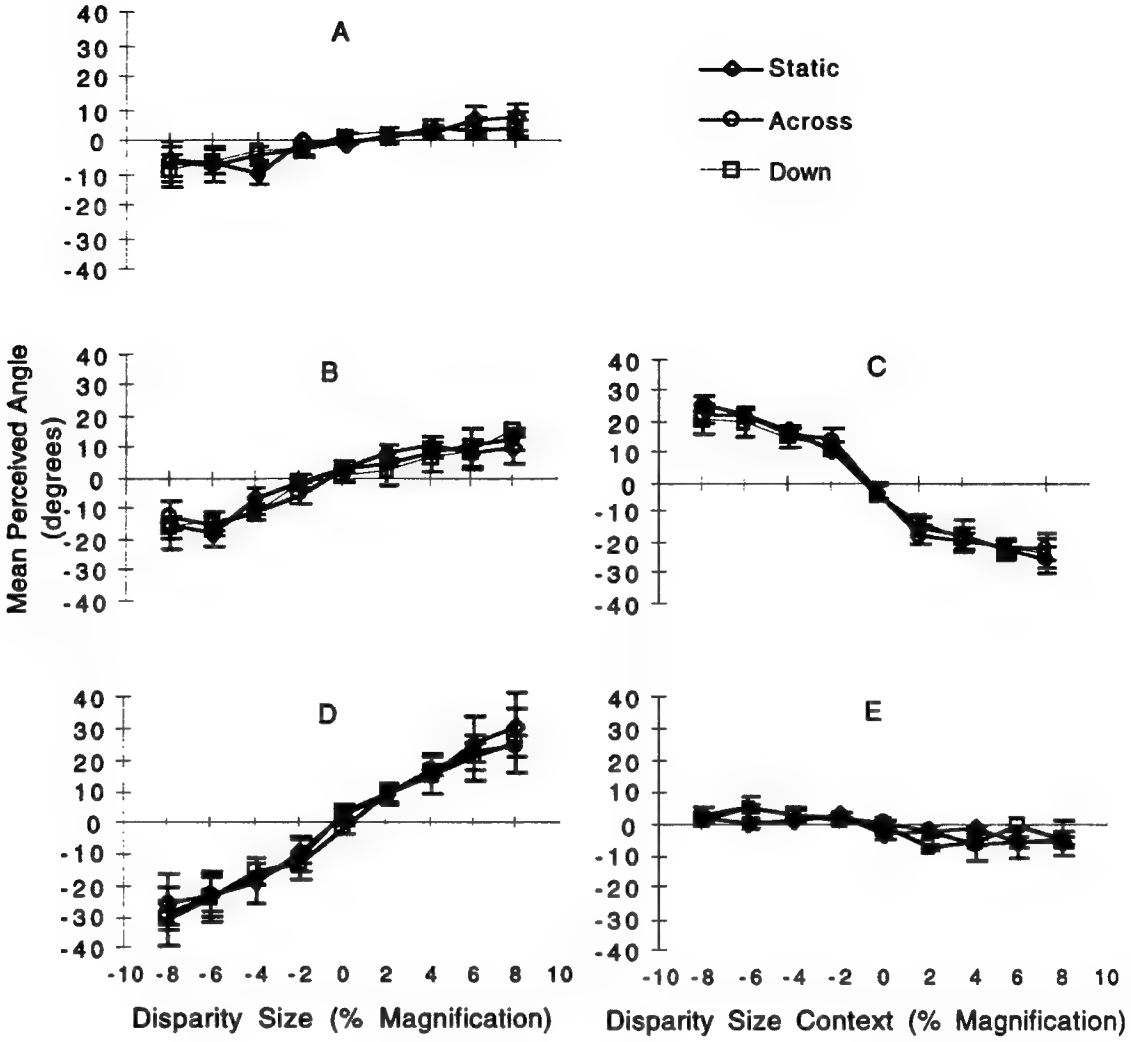
C




A
B
C
D
E
Static
Across
Down
Theoritical
Mean Perceived Angle (degrees)
40
30
20
10
0
-10
-20
-30
-40
-10 -8 -6 -4 -2 0 2 4 6 8 10
Disparity Size (% Magnification)
Disparity Size Context (% Magnification)


A
B
C
D
E
Static
Across
Down
Theoritical
Mean Perceived Angle (degrees)
Disparity Size (% Magnification)
Disparity Size Context (% Magnification)


A
B
C
D
E
Static
Across
Down
Mean Perceived Angle (degrees)
40
30
20
10
0
-10
-20
-30
-40
-10 -8 -6 -4 -2 0 2 4 6 8 10
Disparity Size (% Magnification)
Disparity Size Context (% Magnification)


A
B
C
D
E
Static
Across
Down
Mean Perceived Angle (degrees)
40
30
20
10
0
-10
-20
-30
-40
-10 -8 -6 -4 -2 0 2 4 6 8 10
Disparity Size (% Magnification)
Disparity Size Context (% Magnification)

THE ROLE OF MULTI-MODAL ADAPTIVE INTERFACES IN PROVIDING FLIGHT PATH GUIDANCE

Brian W. Moroney
Graduate Student
Department of Psychology

University of Cincinnati
Mail Location 0376
Cincinnati, OH 45221-0376

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

September 1997

THE ROLE OF MULTI-MODAL ADAPTIVE INTERFACES IN PROVIDING FLIGHT PATH GUIDANCE

Brian W. Moroney
Department of Psychology
University of Cincinnati

Abstract

Advances in weapons capabilities and reduced target detectability are requiring fighter pilots to make more time-critical decisions in a shorter period of time (Hettinger & Haas, 1993). Operating in such an environment leads to reduced situational awareness, increased workload, and potentially life-threatening errors. One means proposed to help alleviate the excessive demands imposed on pilots is through the use of real time *adaptive interfaces* which are continuously modified based on the internal state of the pilot, the behavioral pattern of the pilot, and environmental-external events (Hettinger, Cress, Brickman, & Haas, 1996). Simply stated, the goal of adaptive interfaces is to provide the right information, in the right format, at the right time (Bennett, 1997). However, prior to achieving this goal there are a number of issues which must be addressed (Hettinger et al., 1996). First, the conditions under which an adaptive interface will facilitate performance must be identified. Second, rules to guide the appropriate modification of adaptive interfaces must be generated. More simply, the "when" and "how" of adaptive interfaces must be determined.

The proposed study is designed to address both of these issues within the context of a tactical flight environment. The research described in this proposal will examine the performance implications afforded by the use of multi-modal adaptive interfaces during the recovery-to-pathway segment of a precision navigation flight task. This research will examine the implementation of visual and auditory interfaces designed to aid the pilot in returning to the correct course following the completion of maneuvers performed while evading a surface-to-air missile. The results of this research will have implications for the development of future adaptive interfaces. Specifically, this study aims to identify the conditions under which multi-modal interfaces should be employed and the types of adaptations which will prove beneficial.

THE ROLE OF MULTI-MODAL ADAPTIVE INTERFACES IN PROVIDING FLIGHT PATH GUIDANCE

Brian W. Moroney

Introduction

Advances in weapons capabilities and reduced target detectability are requiring fighter pilots to make more time-critical decisions in a shorter period of time (Hettinger & Haas, 1993). Operating in such an environment leads to reduced situational awareness, increased workload, and potentially life-threatening errors. To combat these factors, within the tactical air environment, research at the Synthesized Immersion Research Environment (SIRE) has progressed in two phases. The first phase involved the development and evaluation of multi-sensory *alternate* display and control technologies, while the second phase involves the development of guidelines for the implementation of such interfaces. Alternate interfaces evaluated during the first phase include combined visual-audition displays (Cunningham, Nelson, Hettinger, Russell, & Haas, 1995), direct vestibular displays (Cress, Hettinger, Cunningham, Riccio, McMillan, & Haas, 1996), brain-body actuated control (Nelson et al., 1996) and haptically-augmented control sticks (Brickman, Hettinger, Roe, Lu, Repperger, & Haas, 1996). The primary goal of this research is to identify non-traditional interface technologies that provide the pilot with salient task information, resulting in improvements in pilot-aircraft performance (Haas, 1996).

The current focus of the SIRE research team is to develop rules to guide the implementation of the more promising alternate interface technologies. The major goal for the proposed research is to identify the conditions under which specific interface configurations will facilitate performance (Hettinger et al., 1996). Specifically, what types of adaptations, or modifications to an interface, will help the pilot deal with the demands of the current task? In many respects this research is novel, yet the questions that must be addressed are common to many current human factors issues. From a broad perspective, this research can be viewed as an attempt to determine the most efficient means for capitalizing upon the differing capabilities of the human and machine. In the past, the distribution of tasks was relatively straightforward; however, the rapidly increasing computational abilities of computers have blurred the traditional characterizing features of humans and machines. Indeed, computers can now perform many of the tasks previously performed by humans.

Increases in computational ability are typically associated with an increase in overall system productivity. However, to make this assumption when dealing with a human-machine system would be a mistake. It is true that if the pilot and machine are utilized in a fashion which exploits the capabilities of each entity, overall system performance should increase; however, this may not always be the case. Overall system performance will increase *only* if the human and machine function in a synergistic manner. The failure to acknowledge, and exploit, the differing capabilities of the human operator and the machine is likely to result in the introduction of inefficient and potentially ineffective systems. This distinction is particularly relevant in the flight domain. Due to the high demands and rapid pace of the dynamic tactical flight environment, it is essential to utilize each component (human and machine) as efficiently and effectively as possible.

The research described in this proposal will examine the performance implications afforded by the use of multi-modal adaptive interfaces during the recovery-to-pathway segment of a precision navigation flight task. This research will examine the implementation of visual and auditory interfaces designed to aid the pilot in returning to the correct course following the completion of maneuvers performed while evading a surface-to-air missile. The results of this research will have direct implications for the development of future adaptive interfaces. Specifically, this study aims to identify the conditions under which multi-modal interfaces should be employed and the types of adaptations which will prove beneficial.

Background

Limitations of Functional Allocation

The traditional approach to alleviating the excessive demands imposed on pilots is to allocate tasks, or a portion of a task, previously performed by a human operator to a machine. While this approach has been successfully implemented in a limited number of cases (e.g., autopilot), it has also resulted in the introduction of new, and perhaps more severe, types of errors (Weiner & Curry, 1980; Woods & Cook, 1991). Specific errors include loss of situational awareness, loss of system awareness, and loss of manual skills (Norman, Billings, Nadel, Palmer, Weiner, & Woods, 1988). Typically, these errors are associated with attempts to implement "static", or fixed automation systems in which the allocation of tasks remains constant in all situations (NADC, 1989).

The errors associated with attempts to implement such systems can be attributed, in part, to the technology-centered design approach utilized during system development. In this case, the function of the system is based on available technological resources and not an evaluation of the role of the operator within the system. This approach results in the development of systems which take the operator "out of the loop", or, as described by Sheridan (1970), shift the role of the human operator from that of an active controller to that of a supervisor who serves in a fail-safe capacity in the event of system malfunction.

As discussed by Parasuraman (1987), limiting operators to the role of a passive monitor seems to be a task for which humans are ill-suited. Earlier work by Wickens and Kessel (1981), which evaluated participant's abilities to monitor a tracking task for changes in the dynamics of the task, provides evidence for this view. In one condition, participants were required simply to monitor the tracking task, while in the remaining condition, they were required to monitor and manually control the same task. The authors reported that response times to detect a change in system dynamics were slower for participants in the monitoring only condition. In a similar study, Bortolussi and Vidulich (1989) reported that helicopter pilots utilizing an automated hover system had slower response times to unexpected events occurring during the hover maneuver. The findings of these studies affirm the existence of drawbacks associated with the current approach to functional allocation.

In an attempt to combat these apparent costs, an alternative approach has been proposed, that of utilizing *adaptive interfaces* designed to preserve the human operator as an active participant in system operations. The development of adaptive interfaces is based on a reevaluation of the traditional view of functional allocation in which tasks are assigned in an all-or-none fashion. The proposed approach utilizes a human-centered design

viewpoint, in which the aim is to develop systems which complement the perceptual, cognitive, and motor control capabilities of the operator.

Adaptive Interfaces

Simply stated, properly designed adaptive interfaces provide the right information, in the right format, at the right time (Bennett, 1997). Through the implementation of such interfaces, it is anticipated that the problems associated with the traditional functional allocation approach can be avoided. Unlike existing human-computer systems that typically utilize a functional allocation policy, adaptive interfaces do not alter the control of the task. Instead, the adaptive interface approach is concerned with *information allocation*, that is, providing the pilot with the information necessary to deal with the demands of the current task.

To enable allocation of the appropriate information, adaptive interfaces are continuously modified based on a series of interdependent "triggers." Potential triggers include internal-physiological measures such as EEG, heart rate, respiration rate, and eye blink rate, behavioral events such as eye/head movements and control stick inputs, and external-environmental events such as the appearance of a target, or hydraulic system failure (Brickman et al., 1995; Hettinger et al., 1996). Through the weighted combination of these three interdependent triggers, an adaptive logic algorithm can be developed to modify the interface configuration, and thus provide the appropriate information.

Research efforts thus far have focused on evaluating individual components of the adaptive logic algorithm. For example, to evaluate the utility of behavioral events such as control stick inputs, Brickman et al. (1995) examined the implementation of a force-reflecting haptically-augmented aircraft control stick in both calm and turbulent flight conditions. The haptic stick provided information about the lateral deviation of an aircraft during an ILS approach to a runway. When the aircraft remained aligned with the runway the control stick operated normally; however, if the aircraft deviated laterally from a predetermined flight path, the stick was modified to provide increased resistance against control inputs which would further increase this deviation. In contrast, control inputs guiding the aircraft toward the correct lateral alignment were met with normal (i.e., less) resistance. The authors reported that touchdowns under turbulent flight conditions (measured in terms of both lateral and longitudinal RMS error) were more accurate with the haptic stick as compared to the normal stick. Under calm flight conditions there were no performance differences between the two control stick conditions. Thus, Brickman and his colleagues identified a condition (turbulence) in which an adaptive interface (the haptic stick) facilitated performance (more accurate landings).

A more recent study, conducted by Bennett (1997), evaluated the utility of a haptic stick and a modified flight director during a low-level precision navigation flight task. As in the previous study, the haptic stick provided course deviation information through changes in the stick dynamics (i.e., increased resistance for inputs guiding the aircraft off course). In this case, however, the information provided by the haptic stick was supplemented through the use of a modified HUD (MHUD). The MHUD, as shown in figure 1b, consisted of a standard heads-up-display (HUD; figure 1a) with the addition of a flight director box. The pilot's task was to keep the waterline icon (w) inside the flight director box that moved based on the position of the aircraft relative to the objective flight path. Maintaining the waterline in the box enabled the pilot to maintain the correct altitude and heading.

In this study Bennett employed a number of conditions; pilots performed the flight task using a standard HUD (figure 1a), the MHUD in a constantly on mode, in which it was presented for the entire flight, or in an adaptive mode, in which the MHUD was presented only when the pilot deviated from the correct course. The results indicated that flight task performance (i.e., the ability to maintain the correct flight path) was superior while using the MHUD, in either the constant or adaptive modes, as compared to the standard HUD. This improved flight performance using the MHUD was not surprising, as previous studies have shown that flight directors can improve navigational task performance (Reising, Liggett, Solz, & Hartsock, 1995). However, the lack of performance differences for the constant and adaptive modes of the MHUD was not anticipated.

The failure to find such implementation mode differences, as discussed by Bennett (1997), may be due to the nature of the task. In this study, pilots only performed the navigational flight task. Thus, some of the additional tasks that pilots perform in a typical flight, such as searching the visual field for targets and monitoring audio communications, which may influence the efficacy of an adaptive interface were missing. The current study aims to explore this possibility through the implementation of additional visual and auditory tasks intended to mimic the modality-specific demands imposed on pilots. This manipulation will afford an examination of the performance implications of modality-specific demand on visual and auditory interfaces designed to assist the pilot return to the correct flight path. One of the goals for the multi-modal interfaces under examination is to provide pilots with the information necessary to navigate following the highly disorienting maneuvers associated with evading a surface-to-air missile, as discussed in the following section.

Navigation and Situational Awareness

Situational awareness (SA) is a multi-dimensional construct involving the comprehension of current surroundings, the changing situation, mission objectives, the ability to act upon relevant information, and the anticipation of future events (Shrestha, Prince, Baker, & Salas, 1995). Simply stated, it is the ability to "figure out what is going on", so the appropriate course of action can be taken (Adam, 1994). Included in this concept is the ability to navigate, which involves determining current position relative to objective destination. To effectively navigate a modern aircraft the pilot must integrate information from numerous sources. For example, while navigating non-GPS equipped aircraft under instrument flight conditions (i.e., with no external visual referents), current location must be updated by the integration of altitude, heading, airspeed and bearing information obtained from displays in the cockpit. The failure to correctly integrate this information can lead to a loss of SA and potentially fatal results (see Aeronautica Civil of the Republic of Colombia [1995] which describes the crash of American Airlines Flight 965).

Numerous authors have suggested that auditory and visual interfaces may prove beneficial in improving SA (Calhoun, Janson, & Valencia, 1988; Reising, Liggett, Solz, & Hartsock, 1995; Wenzel, Wightman, & Forster, 1988). While no studies have formally examined the use of an auditory interface in improving SA with respect to navigation, McKinley, Erikson, and D'Angelo (1994) used 3-D audio to help pilots localize multiple ground-based targets. The authors reported that pilots performing the inflight localization task indicated that the 3-D audio contributed to their SA. In the case of visual interfaces, Reising, Liggett, Solz, & Hartsock, (1995) developed and

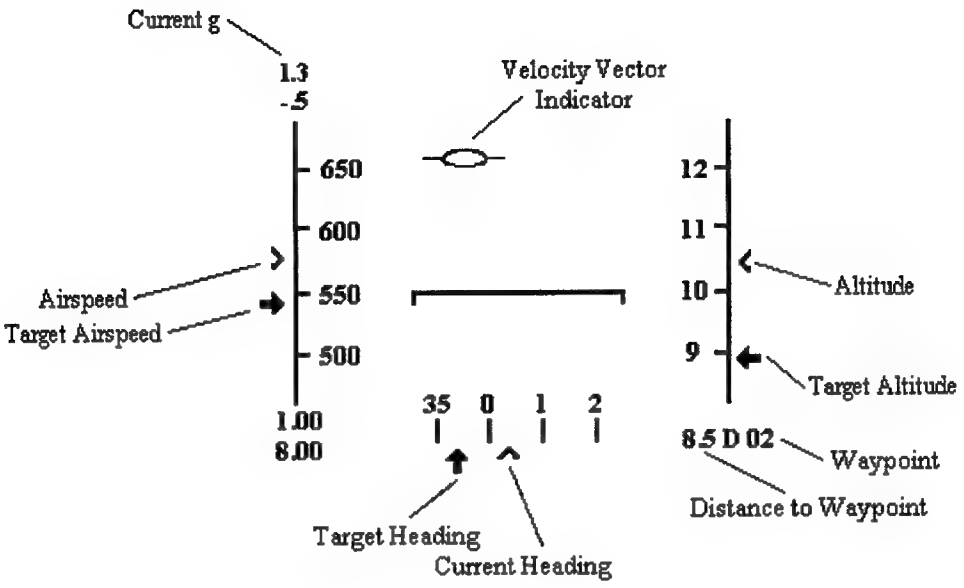


Figure 1a: Standard HUD

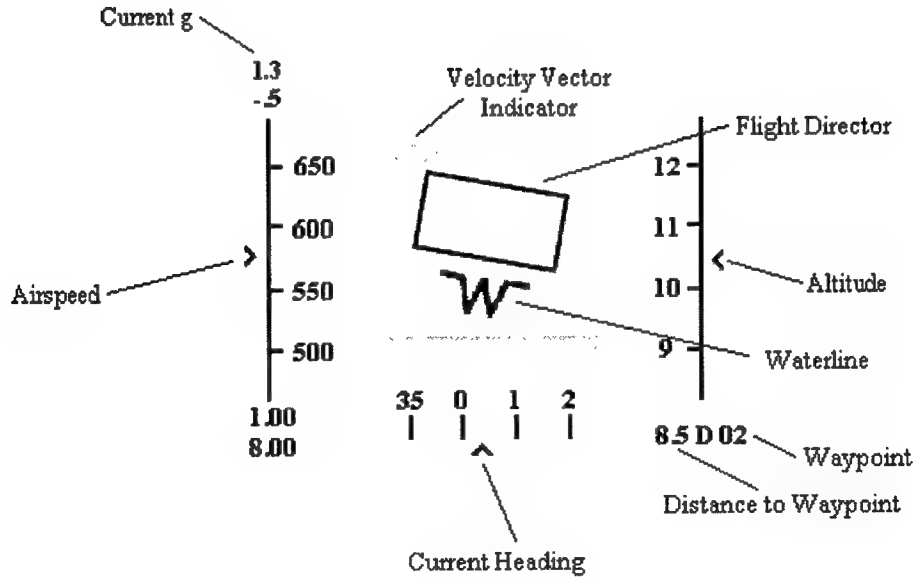


Figure 1b: Modified HUD (MHUD)

evaluated a flight director HUD which utilized a pathway-in-the-sky interface. While the authors reported improved SA, Lintern, Waite, and Talleu (1997), suggest that these results must be interpreted with caution. Lintern et al. suggest that the improvement in SA may be due to the pilot's perception of how well they were following the flight path and not an improvement in the more global level of SA. Nonetheless, the findings of McKinley et al. and Reising et al. suggest that improved visual and auditory interfaces have the potential to improve SA. The types of auditory and visual interfaces to be employed in the current investigation are discussed in the following sections.

Auditory Interfaces

The development of auditory interfaces for use in flight navigation can be traced to the 1930's. The focus of the initial research was on developing auditory interfaces that enabled pilots to control their aircraft under degraded visual conditions (DeFlorez, 1936). Due to advances in computational ability, this topic has seen renewed interest in recent years; however, the research focus has experienced a dramatic shift. Researchers are no longer interested in replacing visual interfaces with auditory interfaces, but rather are concerned with utilizing both visual and auditory interfaces in a synergistic manner. That is, the intent is to design systems that utilize multi-modal interfaces to facilitate overall system performance. For example, a current research program in the SIRE facility is evaluating a multi-modal high off-boresight targeting system. This system utilizes a visual interface for targets located within a specified angle of the pilot's line-of-sight, but switches to an aural interface if the target location exceeds the pre-determined limits (Tannen, 1997).

In addition to a shift in focus, the scope of auditory interface research has also increased. While the initial auditory interfaces were designed to reduce the visual load imposed by environmental sources such as fog (DeFlorez, 1936; Forbes, 1946), the current generation of auditory interfaces is designed to reduce the load imposed by sources both internal and external to the cockpit. As suggested by McKinley, Erikson, and D'Angelo (1994), one source of increased visual load can be attributed to the rising complexity of modern aircraft. In an attempt to combat this visual load, integrated interfaces such as head-up (HUD) and helmet mounted displays (HMD) have been developed. However, these interfaces have proven only partially effective (McKinley et al., 1994). Thus, the development of alternate information-presentation techniques such as auditory interfaces is imperative if the full capability of the next generation of tactical fighters is to be attained.

The use of auditory information to aid pilots perform visually demanding tasks, such as a landing approach under reduced visibility conditions, is not novel. Indeed, the Standard Beam Approach, in which landing deviation information was represented by the acoustic presentation of Morse code letters, was successfully employed prior to World War II (Ellis, Burrows, & Jackson, 1953). In this approach, the letter A ("dit dah") represented deviation in one direction, while the letter N ("dah dit") represented deviation in the opposite direction. If correctly aligned with the runway a steady tone was presented (Lyons, Gillingham, Teas, Ercoline, & Oakley, 1990).

In 1936, DeFlorez proposed a similar system designed to aid pilots navigate from waypoint to waypoint (i.e., the next turning point). In contrast to the Standard Beam Approach, which required an interpretation of the meaning of the letter prior to inputting a corrective action, DeFlorez proposed to utilize the binaural lateralization capabilities of the pilot. The proposed system consisted of a pair of direction finding radio loops positioned slightly

apart, thus enabling a method of triangulation on the broadcasting tower designated as the next waypoint. As long as the aircraft remained aligned with the waypoint, signal intensity remained equal in both ears. However, whenever the aircraft veered off course, the intensity of the sound in the ear closer to the tower increased. While the proposed system is intriguing, it is not known whether it was ever utilized. In addition, its utility is limited in principle by the fact that the sound intensity is based on the location of the aircraft relative to the tower, and not the orientation of the pilot's head.

A more recent study by Loomis, Herbert, and Cicinelli (1990) evaluated a system with a similar approach. In this case, however, virtual sounds were generated using a 3-D audio system which incorporated observer head position. Thus, as the head of the observer moved, the spatial location of the sound remained fixed. The system used in that study was developed as a means of assisting visually impaired individuals to navigate in a natural environment. To evaluate the navigational ability of participants using the 3-D audio system, Loomis et al. generated virtual sounds and instructed participants to walk to the source of the sound. The results indicated that participants, both blind and sighted, were able to navigate accurately using virtual sounds. In addition, by requiring participants to navigate to real sounds, the authors determined that any navigational performance differences for real and virtual sounds were negligible. These results are supported by additional studies which indicate that performance on localization tasks is similar for participants under real and virtual sound presentation modes (Wightman & Kistler, 1989; Wenzel, Arruda, Kistler, & Wightman, 1993). Taken together, these results provide support for the notion that humans can accurately navigate using virtual sound information.

Visual Interfaces

Current cockpit displays have a number of limitations. First, as suggested by Adam (1994), 70-80% of the instrument panel in a tactical aircraft is inflexible. Thus, scenario-specific information, such as weapons targeting and threat information must be condensed into the remaining 20-30% of the available display surface. Based on advances in sensor and data linking technologies, it is anticipated that the problems associated with limited display real estate will increase unless the current approach to interface design is reevaluated. Attempts to improve this situation have led to the second limitation of current displays. To provide the pilot with access to the information necessary to deal with a variety of situations, interface designers have utilized a layered, or mode approach in which the pilot can locate the needed information by selecting the appropriate menu (Sarter & Woods, 1994). While this approach affords access to more information, it also burdens pilots with the task of display management. Thus, pilots must select the correct menu or mode to access the necessary information or achieve the desired system changes. Note that the above assumes that pilots know how to locate the desired menu or select the appropriate mode of operation; they must still perform the steps necessary to activate it. Providing the pilot with the appropriate information in a timely manner is particularly relevant in tactical aircraft, since the typical hostile engagement ranges from approximately 30 seconds to 2 minutes (Adam, 1994).

Current displays are also limited in type. The majority of existing interfaces in tactical aircraft can be classified as status displays (Stokes & Wickens, 1988). That is, they provide information such as airspeed, altitude, and heading that indicate the current status of the aircraft. Thus, to perform a complex task, such as navigating in

unfamiliar territory, the pilot cannot directly perceive his/her position relative to the destination, but must integrate spatial, temporal, and relational information. To combat these limitations, interfaces which directly represent functional properties have been proposed. For example, the functional property "fuel range" consists of the integration of remaining fuel, weight of the aircraft, and wind velocity (Lintern, 1997). Through the direct representation of fuel range, pilots are able to accurately perceive the action-capabilities of the aircraft. This type of representation enables pilots to process information at a lower rule-based level, instead of resorting to the higher, more error-prone, knowledge-based level (Rasmussen, 1983).

An alternate display type is the command display which provides pilots with information about the desired state of the aircraft (Stokes & Wickens, 1988). Instead of requiring the operator to integrate the necessary pieces of information, command displays show the pilot what performance is expected. Prime examples of command displays include guidance displays, such as flight directors, prediction and quickening displays (Lintern, 1997). As relevant to the current investigation, only flight directors will be discussed. The utility of flight directors has been demonstrated for both landing approaches (Lintern, Roscoe, & Sivier, 1990; Reising, Liggett, Solz, & Hartsock, 1995) and navigational flight tasks (Bennett, 1997). In each study, adherence to the optimal flight path was improved as compared to performance using more traditional navigational interfaces. In addition, evidence suggests that flight-director type displays have the potential to improve SA (Reising, Liggett, Solz, & Hartsock, 1995).

Adaptive Interfaces and Modality Specific Demand

As discussed previously, multi-modal adaptive interfaces have the potential of improving overall system performance by providing the right information, in the right format, at the right time (Bennett, 1997). Correctly designed adaptive interfaces will anticipate the information requirements of the pilot and provide the appropriate information in the appropriate format in a timely manner. The potential benefits of a properly designed interface include reduced display- management requirements, improved use of display real estate, improved SA, and a reduction in workload. Conversely, poorly designed adaptive interfaces have the potential to reduce system performance by presenting irrelevant information or removing necessary information (Bennett, 1997).

This logic follows that of Lintern (1997), who suggests that interfaces should be employed in an adaptive fashion only if evidence of a potential conflict with an additional flight task exists. For example, the implementation of an auditory interface under high auditory load conditions, may reduce both primary flight and additional communications monitoring task performance. To predict the performance benefits associated with multi-modal adaptive interfaces, Haas (1996) has suggested that the multiple-resource theory of attention, proposed by Wickens (1984), be employed.

According to Wicken's theory, the human information-processing system consists of "reservoirs" or "pools" of resources from which attention is allocated. These reservoirs can classified based on: (1) stage of processing (early vs. late), (2) code of processing (verbal vs. spatial), and (3) modality of processing (auditory vs. visual). Based on Wicken's theory, it is possible to make predictions regarding the utility of multi-modal interfaces. Thus, employing an adaptive interface in the same modality as the high load condition will degrade dual-task performance. Conversely, the implementation of an adaptive interface in an alternate modality, with respect to that

currently being loaded, should improve dual-task performance. To test this notion, the current investigation will induce modality-specific loads and examine the performance associated with intra- and inter-modality interfaces.

The Present Study

The proposed investigation aims to serve two purposes: first, to serve as an exploratory investigation into the utility of auditory and visual interfaces during the recovery-to-pathway segment of a precision navigation flight task; second, to identify the conditions under which adaptive interfaces will facilitate performance and to generate rules to guide the appropriate modification of adaptive interfaces.

Recovery-to-pathway. One of the primary components of situational awareness involves knowing where you are relative to where you want to be. Thus, situational awareness involves the ability to navigate toward a specific location. Existing aircraft displays are limited in that they require pilots to integrate multiple sources of information to determine current location relative to their objective destination. The failure to correctly integrate this information can lead to a loss of SA. In an attempt to counter the drawbacks associated with existing displays, the current investigation will compare a standard interface configuration (SHUD) with proposed adaptive auditory and visual interfaces designed to assist pilots navigate during the recovery to pathway segment of a simulated flight task.

The auditory interface, or navigational flight beacon (NFB), to be employed utilizes a 3-D audio system that represents the objective destination as a virtual point in space. Evaluations of similar systems have demonstrated that participants can accurately localize virtual sounds in both quiet and noisy conditions, as well as under the highly demanding conditions of low-level flight (McKinley, Erikson & D'Angelo, 1994). The use of an auditory-based navigational interface should capitalize upon the individual's ability to monitor auditory information both inside and outside of the field of view (McKinley et al., 1994). In addition, the NFB should afford navigational information more directly, as compared to existing status displays. Thus, it is anticipated that the use of the NFB will provide the pilots with important navigational information, thus improving recovery to pathway performance and situational awareness.

The modified visual interface (MHUD) to be utilized was developed as a variant of existing flight directors (Bennett, 1997). In this case, however, the pilot is not required to follow a path in the sky, but rather to align the waterline with the flight director box (see figure 1b) which moves based on a comparison of current location to objective destination. Correct alignment of the waterline and flight director box will aid the pilot in recovering to the correct flight path. In essence, the MHUD performs the higher level processing previously performed by the pilot when utilizing a standard HUD. Thus, the use of the MHUD, as with the NFB, should reduce the integration demands imposed on the pilot and improve navigational performance and SA.

The "when" and "how" of adaptive interfaces. As discussed by Hettinger et al. (1996), there are two questions which must be addressed prior to the implementation of adaptive interfaces. First, the conditions must be identified under which an adaptive interface will facilitate performance. Second, rules must be generated to guide the appropriate modification of adaptive interfaces. More simply, the "when" and "how" of adaptive interfaces must be determined.

To examine these questions, both the NFB and MHUD will be employed in an adaptive fashion, such that they are only activated following a deviation from the flight path (e.g., the deviation accrued while evading a SAM). In contrast, the SHUD will be employed such that the same information is presented for the entire flight. The presentation of the NFB and MHUD, in conjunction with additional auditory and visual tasks, will provide an opportunity to examine the performance implications of modality-specific interfaces on additional auditory and visual tasks. As suggested by Lintern (1997), the use of interfaces in an adaptive mode should occur only if the interface may potentially interfere with additional flight tasks. Accordingly, the visual and auditory load tasks are intended to mimic the modality specific demands imposed on pilots.

Based on the multiple-resource theory of attention proposed by Wickens (1992), it is anticipated that dual task performance will be degraded when the adaptive interface is in the same modality as the additional task. Conversely, dual task performance should remain unchanged when the adaptive interface and additional task are in alternate modalities. The use of a dual task paradigm will afford a more suitable evaluation of the performance implications associated with the adaptive interface approach. Indeed, it is conceivable that navigational flight performance may be similar for all interface configurations (SHUD, MHUD, NFB, MHUD + NFB), however there may be differences in the associated dual task performance. Based on the findings of the current investigation it will be possible to generate rules to guide the appropriate modification of future adaptive interfaces.

Experimental Design and Methods

Participants.

Approximately 12 pilots recruited from Wright-Patterson Air Force Base will participate in the current investigation. All participants will be tested for normal or corrected-to-normal vision and hearing. Participants will attend one day of training and three additional days of experimental testing. The training session will consist of one hour of flight instruction. Each experimental session will last approximately two hours.

Experimental Design.

To maximize the number of observations, a two within subjects factorial design will be utilized. The independent variables are interface type (SHUD, MHUD, NFB, MHUD & NFB) and modality load (no load, visual load, and auditory load). Each of these variables is discussed in the following section. The dependent measures to be examined include:

Flight Performance Measures	Behavioral Measures	Subjective Measures
• Pathway recovery time. • Vertical and lateral flight path RMS error. • Timing errors on reaching waypoints. • Crashes	• Dual task performance: - Visual: number of scuds identified - Auditory: number of critical signals detected. • Scanning Patterns	• Situational Awareness measures: CC-Sart • Workload measure: NASA TLX

Apparatus and Procedure.

Task: Pilots will perform a manual terrain-following, terrain-avoidance (TFTA) precision navigation flight task (see Appendix A for additional details) using one of the following interface configurations. All interface configurations, with the exception of the standard configuration, will be employed in an adaptive manner for the entire mission. Approximately midway through the flight a critical event consisting of a simulated SAM attack will occur, thus forcing the pilot to perform evasive maneuvers.

Interfaces: Following the completion of the evasive maneuvers, the pilot will be required to return to the original flight path utilizing one of the following interfaces:

(1) Standard visual interface configuration (SHUD; non-adaptive) - consisting of a baseline HUD (see figure 1a), Radar Warning Receiver (RWR), Attitude Display Indicator (ADI), and Horizontal Situation Display (HSD). The HSD is a moving map type display similar to that currently used in Air Force aircraft (e.g. F-15 and F-16)

(2) Modified visual interface configuration (MHUD; adaptive) - consisting of a modified HUD, RWR, ADI, and HSD. The MHUD requires the pilot to align the waterline with a target box to return to course (see figure 1b).

(3) Standard visual and Navigational Flight Beacon (NFB; adaptive) - consisting of a standard HUD, RWR, ADI, and HSD. In this interface configuration, the pilot may use the NFB and/or the HSD to return to the groundpath. The NFB uses the 3-D audio system to represent a point in space towards which the pilot should orient the aircraft. See Appendix B for the results of a preliminary evaluation of the virtual audio system to be used in the current investigation.

(4) MHUD + NFB (adaptive) - consisting of a standard HUD when on course, a modified HUD when off course, RWR, ADI, and HSD. In this interface configuration the pilot may use the MHUD, NFB, and/or HSD to return to the groundpath.

Modality Load Condition: In an attempt to increase modality-specific demand, a dual-task paradigm will be employed. The load conditions, visual and auditory, will be induced through the introduction of an additional task. The additional task must be performed for the entire flight. The modality-specific conditions are described below.

(1) No load:

- pilots will perform the entire flight with no additional tasks (i.e., pilots will simply fly the course, evade, and return to the course).

(2) Visual load:

- increase visual demand by requiring the pilots to perform a "Scud Hunt" mission during which they must search for ground based targets. When a target is identified, pilots will be required to pull the control stick trigger, thus marking the location of the target.

(3) Auditory load:

- auditory demand will be increased by requiring pilots to monitor a simulated auditory radar return for a change in presentation rate of the repetively presented tones. When a critical signal is detected, pilots will be required to respond by pulling the control stick trigger.

References

Adam, E.C. (1994). Tactical cockpits - The coming revolution. In R.D. Gilson, D.J. Garland, & J.M. Koonce (Eds.), Situational awareness in complex systems. (pp. 101-110). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Bennett, K.B. (1997). Dynamically adaptive interfaces: A preliminary investigation. Unpublished manuscript submitted to the Air Force Office of Scientific Research.

Bortolussi, M.R. & Vidulich, M.A. (1989). The benefits and costs of automation in advanced helicopters: In R.S. Jensen (Ed.) Proceeding of the fifth international symposium on aviation psychology, (pp. 594-599). The Ohio State University: Department of Aviation.

Bronkhorst, A.W., Veltman, J.A., & Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. Human Factors, 38(1), 23-33.

Brickman, B.J., Hettinger, L.J., Roe, M.M., Lu, L., Repperger, D.W., & Haas, M.W. (1996). Haptic specification of environmental events: Implications for the design of adaptive, virtual interfaces. Presented at IEEE Virtual Reality Annual Symposium, Santa Clara, CA.

Calhoun, G.L., Janson, W.P., & Valencia, G. (1988). Effectiveness of three-dimensional auditory directional cues. Proceedings of the 32nd Human Factors Society Annual Meeting, (pp. 68-72). Anaheim, CA: Human Factors and Ergonomics Society.

Cress, J.D., Hettinger, L.J., Cunningham, J.A., Riccio, G.E., McMillan, G.R.M., & Haas, M.W. (1996). An initial evaluation of a direct vestibular display in a virtual environment. Proceedings of the Human Factors Society 40th Annual Meeting, (pp.). Santa Monica, CA: Human Factors Society.

Cunningham, J.A., Nelson, W.T., Hettinger, L.J., Russell, C., & Haas, M.W. (1995). Assessing target detection performance using head position data: A qualitative and quantitative analysis. In Proceedings of the Eighth International Symposium on Aviation Psychology, Columbus, OH.

DeFlorez, L. (1936). True blind flight. Journal of the Aeronautical Sciences, 3, 168-170.

Ellis, W.H.B., Burrows, A., & Jackson, K.F. (1953). Presentation of air speed while deck landing: Comparison of visual and auditory methods. Flying Personnel Research Committee Report 841. R.A.F. Institute of Aviation Medicine.

Forbes, T.W. (1946). Auditory signals for instrument flying. Journal of the Aeronautical Sciences, 13, 255-258.

Haas, M.W. (1996). Understanding attentional mechanisms underlying increases in performance associated with multi-sensory displays. (AL/CFHP - DSN785-8768). WPAFB, OH: Armstrong Laboratory.

Hettinger, L.J., & Haas, M.W. (1993). Virtual envirnoments and the ecological theory of perception.. Unpublished manuscript.

Hettinger, L.J., Cress, J.D., Brickman, B.J., & Haas, M.W. (1996). Adaptive interfaces for advanced airborne crew stations. Proceedings of the 3rd Annual Symposium on Human Interaction with Complex Systems. Dayton, OH.

Lintern, G., Waite, T., & Talleu, D.A. (1997). AGATE cockpit design for efficiency of training. (Technical Report ARL-97-4/NASA-97-2). Hampton, VA: NASA Langley Research Center.

Lintern, G., Roscoe, S.N., & Sivier, J.E. (1990). Display principles, control dynamics, and environmental factors in pilot training and transfer. Human Factors, 32, 299- 317.

Loomis, J.M., Herbert, C., & Cicinelli, J.G. (1990). Active localization of virtual sounds. Journal of the Acoustical Society of America, 88 (4), 1757-1763.

Lyons, T.J., Gillingham, K.K., Teas, D.C., Ercoline, W.R., & Oakley, C. (1990, August). The effects of acoustic orientation cues on instrument flight performance in a flight simulator. Aviation, Space, and Environmental Medicine, 699-706.

McKinley, R.L., Erikson, M.A., & D'Angelo, W.R. (1994, May). 3-dimensional auditory displays: Development, applications, and performance. Aviation, Space, and Environmental Medicine, A31-A38.

NADC (1989). Adaptive function allocation for intelligent cockpits. Warminster, PA: Naval Air Development Center

Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Roe, M.M., Lu, L., Haas, M.W., Dennis, L.B., Pick, H.L., Junker, A., & Berg, C.B. (1996). Brain-body actuated control: assessment of an alternative control technology for virtual environments. Proceedings of the 1996 IMAGE Conference, Scottsdale, AZ.

Norman, S. Billings, C.E., Nadel, D. Palmer, E. Weiner, E.L., & Woods, D.D. (1988). Adaptive automation philosophy: A source document. NASA, Ames Research Center, CA.

Parasuraman, R. (1987). Human-computer monitoring. Human Factors, 29, 695-706.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 257-266.

Reising, J.M., Liggett, K.K., Solz, T.J., & Hartsock, D.C. (1995). A comparison of two head up display formats used to fly curved instrument approaches. Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society. (pp. 1-5). Santa Monica, CA: Human Factors and Ergonimcs Society.

Sarter, N.B. & Woods, D.D. (1994). "How in the world did I ever get in that mode?" Mode error and awareness in supervisory control. Situational awareness in complex systems. (pp. 111-123). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Sheridan, T. (1970). On how often the supervisor should sample. IEEE Transactions on System Science and Cybernetics SSC-6, 140-145.

Shrestha, L.B., Prince, C., Baker, D.P., & Salas, E. (1995). Understanding situation awareness: Concepts, methods, and training. In W.B. Rouse (Ed.), Human/technology interaction in complex systems, Vol. 7. Greenwich, CO: JAI Press LTD.

Stokes, A.F. & Wickens, C.D. (1988). Aviation displays. In E.L. Wiener & D.C. Nagel (Eds.) Human factors in aviation. (pp. 387-431) San Diego, CA: Academic Press.

Tannen, R.S. (1997). Integrating multisensory displays for an adaptive head-coupled target leading interface. Unpublished manuscript submitted to the Air Force Office of Scientific Research.

Weiner, E.L., & Curry, R.E. (1980). Flight-deck automation: Promises and problems. Ergonomics, 23(10), 995-1011.

Wenzel, E.M., Wightman, F.L., & Forster, S.H. (1988). Three-dimensional acoustic information. Proceedings of the 32nd Human Factors Society Annual Meeting, (pp. 86-90). Anaheim, CA: Human Factors and Ergonomics Society.

Wenzel, E.M., Arruda, M., Kistler, D.J., & Wightman, F.L. (1993). Localization using nonindividualized head-related transfer functions, Journal of the Acoustical Society of America, 94, 111-23.

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.) Varieties of attention. New York, NY: Harper Collins.

Wickens, C.D. (1992). Engineering psychology and human performance. New York, NY: Harper Collins.

Wickens, C.D., & Kessel, C. (1981). Failure detection in dynamic systems. In J. Rasmussen and W.B. Rouse (Eds.), Human detection and diagnosis of system failures. New York: Plenum.

Wightman, F.L., & Kistler, D.J. (1989). Headphone simulation of free-field listening. II: Psychological validation, Journal of the Acoustical Society of America, 85, 868-878.

Woods, D.D., & Cook, R.I. (1991). Nosocomial automation: Technology-induced complexity and human performance. In Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics (pp. 1279-1282). New York: IEEE.

Appendix A
Mission Script

Flight Task:

- The pilot will be instructed (prior to the beginning of the study) to perform a manual terrain-following-terrain-avoidance (TFTA) precision navigation flight task with a goal altitude of approx. 200-500 above the ground.

Starting Point:

- The pilot will begin the task at an altitude of approximately 500-1000 ft (AGL) and a speed of 500 knots heading towards the first waypoint.
- The pilot will begin with a standard interface configuration (SHUD; non-adaptive) - consisting of a baseline HUD (see figure 1a), RWR, ADI, and HSD.
- Adaptive interfaces will be employed based on position of the aircraft relative to the correct ground track. If the aircraft exceeds a pre-determined boundary, the interface will be turned on.

Segment 1:

- This segment will be approximately 3 minutes long, and will end when the pilot reaches waypoint # 2. It will consist of manual TFTA.
- When the pilot is on the correct path (i.e. going from waypoint 1 to waypoint 2) the SHUD configuration, utilizing bugs and a HSD, will be employed in all conditions.
- Adaptive interfaces will be employed if the pilot deviates from the path at any point (except evasive maneuvers) during the entire flight.

Segment 2:

- On path: When the pilot is on the correct path (i.e. going from waypoint 1 to waypoint 2) the SHUD will be employed in all conditions.

Segment 3:

- Missile Launch: consists of the launching of a SAM. Pilots will be presented with missile location information on their radar warning receiver (RWR) along with an auditory warning. SAM launches will occur within the first 60 seconds of segment 3 (approximately 4 to 5 minutes into the flight).
- Evasive Maneuvers: consisting of a "Horizontal-S" maneuver will be performed by the pilot to evade the SAM. Based on the duration of the missile flight the evasive maneuvers will last approx. 30 to 60 seconds.

Segment 4:

- Recovery to Pathway: Following the conclusion of evasive maneuvers an adaptive interface (i.e. Modified HUD, Navigational Flight Beacon alone, or SHUD & NFB) will be implemented. The adaptive interface will remain on until the pilot has returned to the flight path. Once back on course, the SHUD will be turned back on.

Segment 5:

- Back on path:

Conditions:

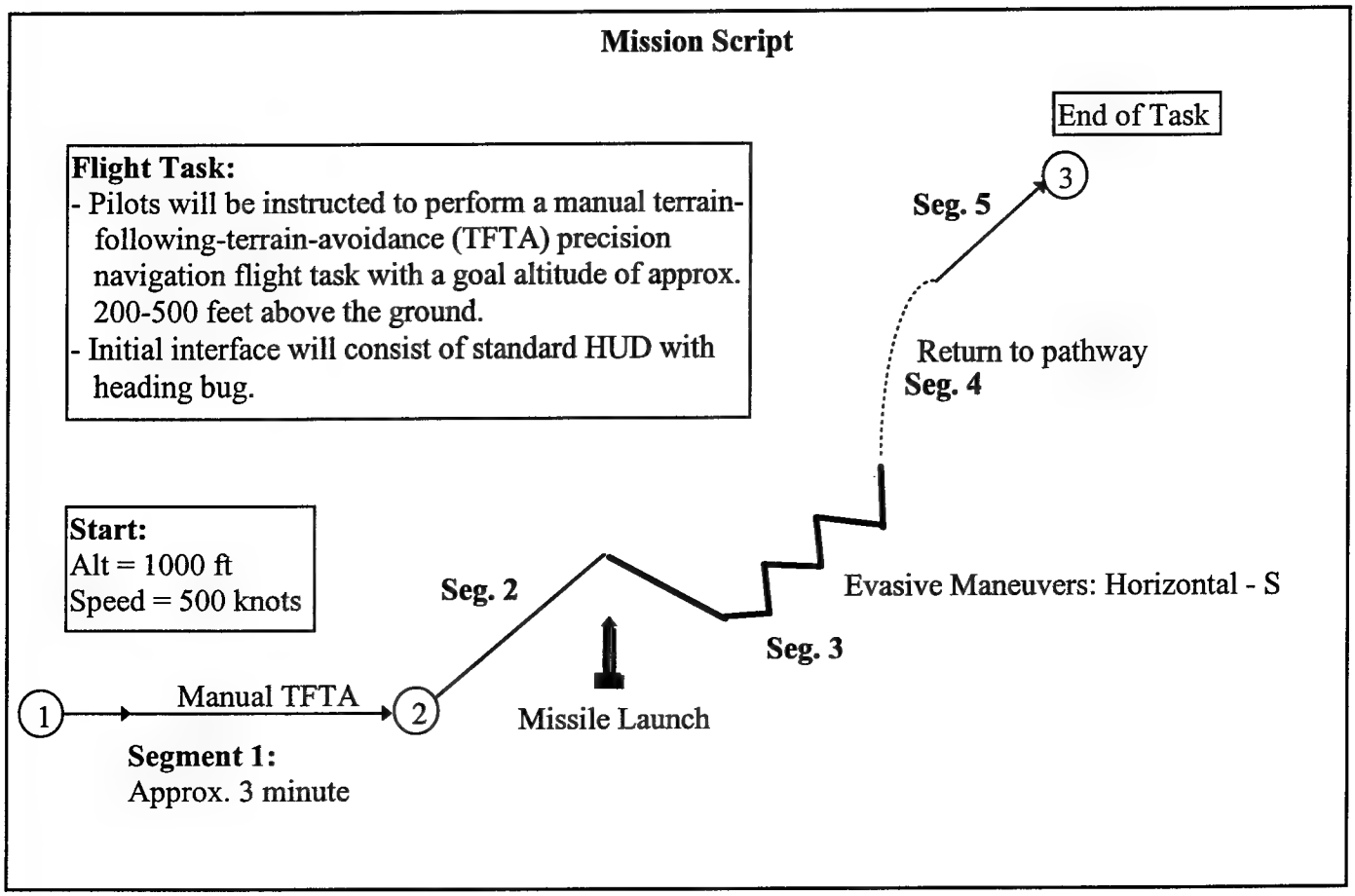
- Pilots will utilize each interface (SHUD, MHUD, NFB, MHUD + NFB) under each of the following conditions.
- The load condition tasks will be performed during the entire flight.

(1) No load: Pilots will simply fly the course.

(2) Visual load: In an attempt to visually "load" the pilot a dual-task paradigm will be employed. The high visual load will be induced through the implementation of a visual search task. While performing the low level flight task, pilots will be required to search the visual field for ground based targets (scud launchers). Once a target is identified pilots will be required to pull the trigger to mark the location of the threat. Pilots will not be required to turn towards the target or lock on the target, only mark the location.

(3) Auditory load: auditory demand will be increased by requiring pilots to monitor a simulated auditory radar return for a change in presentation rate of the repetively presented tones. When a critical signal is detected,

pilots will be required to respond by pulling the control stick trigger. The series of repetitive tones will presented in conjunction with continuous background chatter.



Appendix B
Convolvotron Resolution Preliminary Results

Summary:

Preliminary results indicate that the localization resolution (mean deviation error = 5.76 degrees) errors using the convolvotron are similar to those found in previous virtual auditory investigations (McKinley et al., 1994). It is important to note that all sounds were presented at 0 degrees elevation. Thus, localization errors are likely to increase as elevation deviations are introduced.

Method:

Participants (N = 5) were seated in the cockpit and asked to localize bursts of white noise presented at 30 random azimuth locations by pointing their head at the perceived sound source. Sounds were always presented at 0 degrees elevation. Participants were instructed to inform the experimenter, through verbal communication, once they believed they had localized the sound. Actual sound location, azimuth head position, elevation head position, and deviation information were then recorded by the experimenter. Direction of localization errors (+ or -) were computed following the conclusion of the experimental session. A generic HRTF was used for all participants.

Results:

Mixed results. Deviation angles (|generated sound location - perceived location|) ranged from .01 to 19 degrees with a mean deviation angle of 5.76 degrees for 5 participants. These results fall within the bounds suggested by previous virtual auditory studies (McKinley et al., 1994). However, the distribution of localization errors appears to be skewed to the left (see Figures 1- 5 attached). Of the 150 measures taken (30 per Ss), 109 were to the left of the actual sound location, while only 41 were to the right. In addition, the errors to the left were of much larger magnitude (mean = -6.18) than the errors to the right (mean = 2.94).

To explore the possibility that a hardware / software problem was behind this shift to the left, an oscilloscope was used to evaluate the output produced by the convolvotron following the introduction of a pure tone (created by function generator). Results indicated that the output produced by both channels of the convolvotron were in phase and had equal amplitudes. This test was performed at two locations: 1. directly out of the convolvotron, and 2. at the last point prior to the headphones (wires leading into headphones).

It is believed that the shift to the left was due to the use of a generic HRTF. Further investigations will explore this possibility.

Next Step:

Conduct a similar evaluation with the following independent variables:

- Ambient Noise (baseline Vs. all displays running)
- HRTF (generic Vs. individual)

Figures 1- 5

- Plots of azimuth and elevation localization errors for each participant
- X axis is azimuth. Y axis is elevation.
- Actual sound location is 0,0.
- Notice that figure 1 has a fairly normal distribution (i.e. 15 + and 15 -), while the remaining figures have a negative skew.

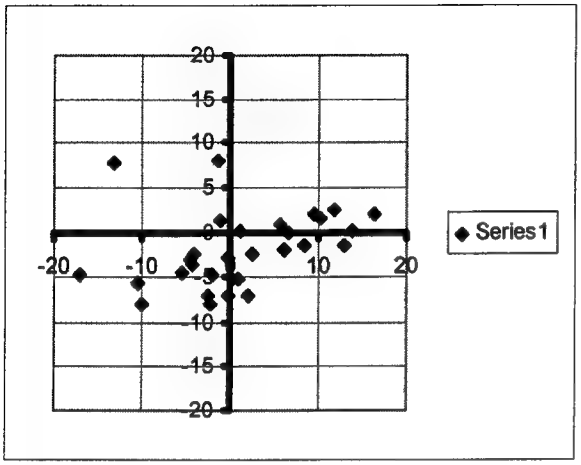


Figure 1

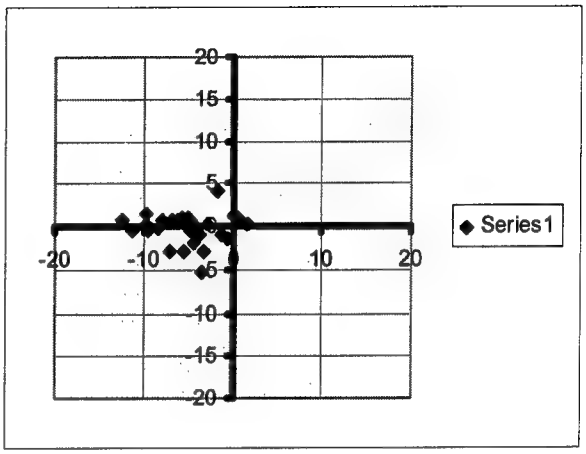


Figure 2

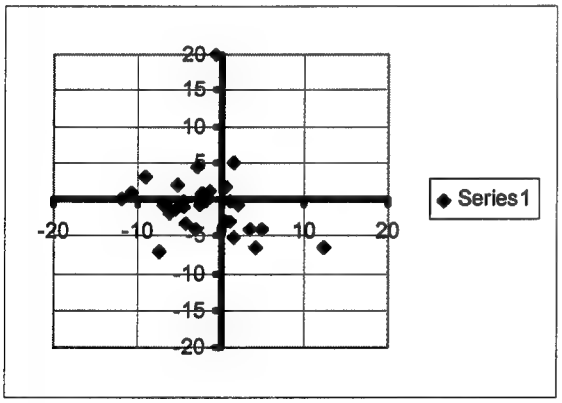


Figure 3

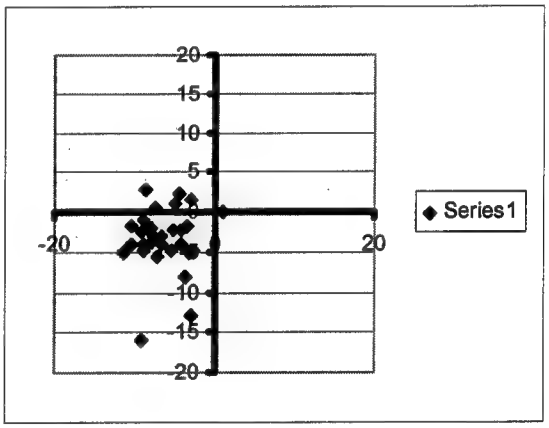
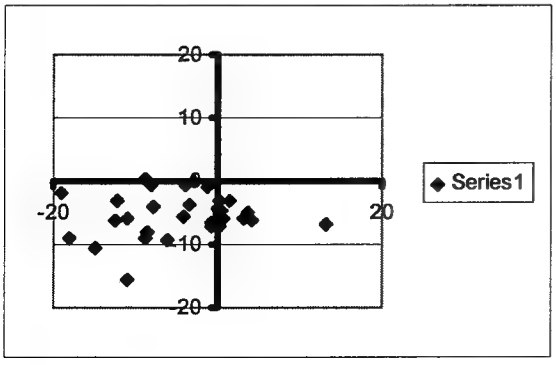


Figure 4



DESORPTION AND BIODEGRADATION OF DINITROTOLUENES IN AGED SOILS

Randy J. Mueller
Graduate Student
Environmental Engineering Program

University of Connecticut
241 Glenbrook Road
Storrs, CT 06269

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

Desorption and Aerobic Biodegradation of Dinitrotoluenes in Aged Soils

Randy J. Mueller
Graduate Student
Environmental Engineering Program
University of Connecticut

Abstract

A study on the desorption kinetics and the biodegradation of dinitrotoluenes in aged soils was performed. Aged clayey soil contaminated with 2,4-DNT, 2,6-DNT, and 2,4,6-TNT was obtained from the Volunteer Army Ammunitions Plant in Chattanooga, Tennessee. The average extractable concentrations of 2,4-DNT, 2,6-DNT, and 2,4,6-TNT were 0.278 (±.004), 0.079 (±.002), and 0.297 (±.004) mg/g, respectively. The organic carbon content (f_{oc}) was 0.70 (± 0.09)%. Using Tenax beads as an infinite sorptive sink, sequential desorption experiments were performed to measure the long-term desorption of DNT and TNT from the soil: 76% and 89% of the extractable nitrotoluenes were desorbed from the soil in one and three days, respectively. Biodegradation of the DNTs was measured in factorial soil slurry experiments. 2,4-DNT was readily available for degradation by the indigenous microorganisms. Complete mineralization, as evidenced by near stoichiometric NO_2^--N release, was achieved within 5 days. Addition of a known 2,4-DNT and 2,6-DNT degrading bacterial strain did not show any benefit. Little mineralization of 2,6-DNT was observed. Additional experiments were performed to evaluate the toxicity or inhibition of 2,6-DNT and TNT towards 2,6-DNT mineralization. The addition of either Tenax beads to reduce the aqueous concentration of 2,6-DNT and TNT or an induced 2,6-DNT degrader (JS922) to the soil slurries did not stimulate 2,6-DNT biodegradation. In separate experiments, it could not be confirmed that TNT at 25 mg/L negatively impacted 2,6-DNT mineralization by JS922. The above work indicates that rapid desorption of nitrotoluenes from the clayey soil tested makes compounds such as 2,4-DNT and 2,6-DNT readily available for biodegradation; that indigenous microorganisms can readily mineralize 2,4-DNT; while stimulation of 2,6-DNT mineralization appears difficult.

DESORPTION AND BIODEGRADATION OF NITROTOLUENES FROM AGED SOILS

by

Randy J. Mueller

INTRODUCTION

Significant contamination of nitrotoluenes in soil and groundwater around the world has prompted an interest in their cleanup. Because nitrotoluenes are toxic, mutagenic and/or potentially carcinogenic they pose a human health risk (6). Although 2,4-dinitrotoluene and 2,6-dinitrotoluene contamination is historically of military origin resulting from TNT manufacturing, they also occur in industry as precursors in polyurethane production. As a result, millions of pounds of DNT are released into the environment each year (6).

Both anaerobic and aerobic biotransformation of DNTs has been reported (12). Anaerobic biotransformation typically is the result of an unspecific reduction of the nitro groups to nitroso, hydroxylamino, and eventually amine groups (12). The potential accumulation of the nitroso and hydroxylamino intermediates, which are more toxic than the parent compounds, plagues anaerobic processes for DNT biotransformation. Many pure cultures have been isolated that can completely mineralize both 2,4 and 2,6-DNT as a sole carbon and energy source. Aerobic biodegradation of DNT involves oxygenolytic denitration pathways, releasing nitro groups as NO_2^- (12, 13).

The overall rate of biodegradation of a compound is determined by its bioavailability (10). Contaminants released in soils can become less bioavailable by sorption into soil organic matter and micropore diffusion into soil particles preventing contact between the contaminant and microorganisms. As the soil ages, the pollutants can diffuse deeper into the soil particle or become covalently bound with the soil organic matter. As a result, aging can further decrease bioavailability.

There were three main objectives to this research. The first objective was to measure the desorption kinetics of DNT in aged contaminated soils in sequential desorption experiments. Second, DNT mineralization kinetics in aged contaminated soil were measured in factorial soil slurry experiments. Finally, the toxicity/inhibition effects of 2,6-DNT and TNT on the biodegradation of 2,6-DNT were investigated.

MATERIALS AND METHODS

Bacterial Inoculum

JS922, a bacterial strain known to degrade both 2,4-DNT and 2,6-DNT, was used for all experiments. It was pregrown at 30°C in shake flasks containing a mineral medium supplemented with 2,4-DNT or 2,6-DNT. The medium contained XAD-7 Tenax beads (Sigma, St. Louis, MO) at 10 g/L to maintain a sub-toxic, aqueous concentration of DNTs. The mineral salts medium, termed MM-C, was a 1/100 dilution of Stanier's mineral salts medium (14) but free of inorganic nitrogen, and provided macronutrients and 0.5 mM pH buffering capacity. The total ionic strength of the mineral salts medium was 3.3 mM. The addition of 50 mg/L or 200 mg/L of $HgCl_2$ increased the ionic strength to 3.5 and 4.3 mM, respectively.

HPLC Analysis

In all experiments performed, high performance liquid chromatography (HPLC) separation followed by a UV-detection at 540 nm was used to measure DNT concentrations. The method used a C6-hexyl column (Spherisorb, Alltech, Deerfield, IL) and a solvent system consisting of 70% deionized water and 30% methanol. Confirmation of HPLC peaks during initial analysis of soils was performed using a second HPLC method using a CN column (Zorbax, Hewlett Packard) and a solvent system consisting of 30% acetonitrile and 70% deionized water (7).

Nitrate Analysis

Samples were centrifuged at 14,000 rpm for 5 min. and the supernatant was analyzed with a Dionex DX-300 Series Chromatography System equipped with a Dionex AS11 column and CDM-2 conductivity detector (Dionex, Sunnyvale, CA). The eluent was 19 mM NaOH at a flow rate of 0.65 ml/min.

Nitrite Analysis

Nitrite analysis was performed by a modified standard method (4). A 0.7 ml sample was centrifuged at 14,000 rpm for 5 min. The supernatant was analyzed colorimetrically by mixing 1 ml of deionized water and 0.1 ml of sulfanilamide with 0.1 ml of sample. The solution was allowed to react for 5 min. before 0.1 ml of N-(napthyl)diethylenediamine was added. Samples were vortexed and allowed to react for 15 min. prior to reading their absorbance at 543 nm using a UV-spectrophotometer.

Soil Treatment

Three different soil samples were received from the Volunteer Army Ammunitions Plant in Chattanooga, Tennessee. The soils were expected to contain large quantities of 2,4-DNT, 2,6-DNT, and 2,4,6-TNT. The concentration of total extractable DNTs and TNT were measured according to established protocols(7). Two gram aliquots of soil (air dried and homogenized with mortar and pestle to pass through a 30 mesh sieve) were measured in a Teflon® lined vial containing 10 ml of methanol as the extracting agent. The vials were mixed by vortexing for one min. and were shaken at 200 rpm overnight at 30°C. The vials were removed and allowed to settle for 30 min.. A 5 ml aliquot of the sample combined with 5 ml of $CaCl_2$ (5 g/L) was centrifuged at 740 g for 5 min. to separate the soil fines. The supernatant fraction was analyzed by HPLC for total extractable nitrotoluenes.

Cell enumeration

Cells were dislodged from soil by a modification of a published protocol (2). A 2 g aliquot of wet soil was mixed with 20 ml of a solution consisting of 0.5% sodium pyrophosphate (v/v) and 0.05% polyvinyl pyrrolidone at pH 7 and kept in a 50 ml polyprolylenene centrifuge. This mixture was intermittently mixed by vortexing for 30 min. Then 20 ml of $CaCl_2$ (5 g/L) was added and the suspension was centrifuged at 750 g for 5 min. @ 15°C. The soil free supernatant was transferred to a clean tube and used for MPN or AODC analysis. For enumeration of microbial populations from soil slurry experiments, 1 volume of soil slurry was mixed with 2 volumes of the sodiumpyrophosphate/polyvinyl pyrrolidone solution. This mixture was vigorously shaken by intermittent vortexing before MPN or AODC analysis.

AODC analysis was done according to the protocol of Nishino (9) with cells preserved with acidified Lugol's solution. A microwell MPN protocol was developed to enumerate DNT degrading population and total heterotrophic population. The mineral salts medium used was derived from the Stanier's mineral salts medium (14) with the absence of an inorganic nitrogen source, phosphate buffer added at 0.5 mM and all other compounds at 1% of the full medium strength. Individual DNT isomers were added at 10 mg/L. The final pH was between 7.0 and 7.2. Enumeration of the total heterotrophic fraction used the non-specific R2A medium (11). Ten-fold serial dilutions were made using a 96 microwell plate. Individual wells contained 90 µl of growth medium, and 10 µl of cell suspensions were transferred by means of a multichannel pipetter. Six replicates per dilution were performed. Plates were incubated at 30°C for at least 7 days. After incubation, nitrite formation in each well of the DNT plates was quantified by adding 10 µl sulfanilamide solution and 10 µl N-(1-naphthyl)ethylenediamine solution to each well . Color intensity was measured at 562 nm with a EL340 Biokinetics Reader (BioTek Instruments) plates after approx. 15 to 30 min. vs control wells that contained no cell suspension. Cell activity in the R2A plates was measured by addition of 10 µl of 0.5% INT [0.25 g INT in 50 ml of 10mM phosphate buffer]. Formazan formation after several hours incubation was measured at 562 nm vs control wells. A microcomputer program was used to convert the raw data to MPN values (8).

Synthetic Sorptive Matrix

The effect of sorption on 2,4-DNT and 2,6-DNT biodegradation was evaluated using a synthetic sorptive matrix. A 25 mg/L solution of 2,4-DNT and 2,6-DNT in MM-C was added to a series of flasks that contained 0, 0.1, or 1 g wet weight of XAD-4 Tenax beads (Sigma, St. Louis, MO). rinsed with MeOH and H_2O. Flasks were equilibrated for 24 hours prior to addition of a 1 ml inoculum of JS922. Subsequently, flasks were incubated at 30°C with shaking at 200 rpm and mineralization was monitored by periodic sampling and analysis of NO_2^--N and DNT concentration.

Long-Term Desorption Experiments

Two sets of experiments were performed to monitor the long term desorption of nitrotoluenes from the soils. Both experiments used XAD-4 Tenax beads as a sorptive sink with a large capacity for the DNTs and TNT. In the first set of experiments, soil/bead mixtures were incubated for a fixed time period and the beads were separated and replaced by fresh beads and a fresh aqueous phase (3). In the second type of experiment, no sequential addition of beads occurred (2).

One gram of soil was placed in Teflon® lined vials with 0.25 grams of XAD-4 Tenax beads and 10 ml of a MM-C containing 50 mg/L of $HgCl_2$ (to inhibit biological activity). Vials were incubated on a shaker table (200 rpm) at 30°C. After one day of incubation, the beads from quadruplicate vials were collected. By adding 1 gram of KCl, the density of the aqueous solution was increased causing the beads to float. The beads could then be separated from the aqueous phase using a separatory funnel (Kimax, Fisher, Pittsburgh, PA). Five ml of methanol was used to extract the beads for 24 and 48 hours. Samples from the MeOH extract were analyzed by HPLC. To the remaining soil, 10 ml of fresh MM-C solution and 0.25 g of XAD-4 Tenax beads were added and the vials were reincubated. This procedure was repeated on days 3, 7, 14, 21, and 28. Non-sequential desorption experiments were set up exactly as described for sequential desorption experiments, but duplicate samples were analyzed on days 1, 3, 7, 14, 21 and 28.

Soil Slurry Experiments

Soil slurry experiments were performed in flasks that contained 25 grams of soil B plus 250 ml of MM-C solution. Three different treatment conditions were examined during phase I and phase II of the experiment. All treatments were performed in duplicates. Table 1 lists the treatments used in phase I and in phase II.

Table 1 - Different Treatments in soil Slurry Systems

	Phase I		Phase II	
Flask #	$HgCl_2$ (50 mg/L)	Bioaugmented	XAD-4 (0.1 g)	Bioaugmented
1	Yes	No	Yes	No
2	Yes	No	No	No
3	No	No	Yes	No
4	No	No	No	No
5	No	Yes	Yes	Yes
6	No	Yes	No	Yes

If bioaugmented, flasks were inoculated with 1 ml of JS922 pregrown on 3 mM 2,6-DNT resuspended in 1 ml of MM-C (at a dose of approximately 4.8×10^7 cells/g soil). Flasks were placed on a rotary shaker (200 rpm) at 30°C. Nitrite and nitrate production were monitored colorimetrically and by ion chromatography, respectively. Aqueous TNT and DNT concentrations were analyzed by HPLC. Samples of the soil phase were removed weekly and were MeOH extracted and analyzed for total TNT and DNT.

After several weeks, further treatments were performed (phase II) that involved a second addition of JS922 (at a dose of approximately 3.3×10^8 cells/g soil) or the addition of XAD-4 resin beads to decrease the aqueous concentration of 2,6-DNT and 2,4,6-TNT. Table 1 lists to which flasks XAD-4 resin beads had been added and whether the flasks were re-inoculated with JS922 during phase II.

Toxicity/Inhibition of TNT and 2,6-DNT

The growth of JS922 on 2,6-DNT in the absence and presence of TNT was evaluated. Media containing 2,6-DNT at 10 mg/L and TNT at 25 mg/L or 0 mg/L was prepared in flasks and then diluted to 50 ml with MM-C. The media were then inoculated with JS922 pregrown on a 3 mM solution of 2,6-DNT. The flasks were incubated at 30°C on a shaker table (200 rpm) and nitrite production was monitored.

RESULTS

The soil characteristics of the Volunteer Army Ammunitions Plant soil are presented in Table 2. All three soils were examined for their total MeOH extractable nitrotoluenes. Soil pH was measured by equilibrating 2 g of dried soil with 20 ml of deionized water and measuring the aqueous phase pH. pH of each soil was well within the range permissible for microbial activity. As all soils were derived from the vadose zone, f_{oc} values were within the expected range; soil B had the highest organic carbon content at 0.7%. The mean MeOH extractable 2,4-DNT, 2,6-DNT, and 2,4,6-TNT concentrations for soil B were 0.278 (± 0.004),

0.079 (± 0.004), and 0.297 (± .004) mg NT/g soil, respectively. Because of its high nitrotoluene concentration, soil B was chosen for the experiments.

Table 2 - Characteristics of Soil from the Volunteer Army Ammunitions Plant

	pH	f_{OC} (%)		2,4-DNT mg/g		2,6-DNT mg/g		2,4,6-TNT mg/g	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
A	5.5	0.465	0.058	3.88E-3	1.33E-4	5.03E-3	1.22E-4	ND	
B	6.4	0.701	0.088	2.78E-1	3.84E-3	7.87E-2	1.59E-3	2.97E-1	3.58E-3
C	8.5	0.226	0.005	3.51E-4	2.49E-5	3.51E-4	2.49E-5	ND	

Most probable number enumeration on soil B revealed a resident bacterial population able to denitrate 2,4-DNT ($2.7\ 10^4$/g soil) and 2,6-DNT ($0.38\ 10^3$/g soil). The total cultivable population was $1.0\ 10^6$/g soil and reflected approx. 10% of the total acridine orange direct count.

Figure 1 shows the results of the sequential and non-sequential desorption experiments for 2,4-DNT (Panel A), 2,6-DNT (Panel B), 2,4,6-TNT (Panel C). Rapid partitioning of the desorbable nitrotoluenes was observed from the aqueous phase into the sorptive sink. The sequential desorption data (1, 3, 7, 14, 21, and 28 days) indicate that after 7 days all the desorbable NTs were removed from soil. The non-sequential desorption (3, 7, 14, 21, and 28 days) data confirms this result by showing all the desorbable NTs captured by the sorptive sink after 7 days.

The cumulative mass of desorbable nitrotoluenes is plotted in Figure 2. A first order desorption rate expression was used to fit the cumulative desorbable NT data.

$$M_{d,t} = M_o(1-e^{-kt})$$

M_o = total mass of desorbable NTs (mg)

$M_{d,t}$ = total amount of NTs desorbed at time t (mg)

First order desorption rate coefficients were estimated at 3.56, 1.19, and 1.80 day^{-1} for 2,4-DNT, 2,6-DNT, and 2,4,6-TNT, respectively. These very high rate coefficients confirm the very rapid desorption. After 1 day of incubation 68% (2,4-DNT), 96% (2,6-DNT), and 78% (2,4,6-TNT), of the total MeOH extractable NTs were desorbed from the soil and 83% (2,4-DNT), 100% (2,6-DNT), and 100% (2,4,6-TNT) were desorbed after 28 days.

The kinetics of the biodegradation of 2,4-DNT and 2,6-DNT sorbed onto a synthetic matrix were measured for JS 922. The addition of Tenax beads progressively reduced the aqueous concentration of DNTs causing the rate of DNT biodegradation, as measured by NO_2^- evolution, to decrease.

Factorial soil slurry experiments were performed to measure the bioavailability of DNT for biodegradation and confirm the rapid desorption found in the long-term desorption experiments. The 2,4-DNT degradation and nitrite profiles of phase I are shown in Figure 1. In the abiotic control flasks, $HgCl_2$ at 50 mg/L suppressed microbial activity for up to 10 days. However, after 10 days 2,4-DNT disappearance occurred in these flasks with a near stoichiometric release of nitrite. As 2,4-DNT disappearance and concomitant NO_2^- accumulation was observed in the control soil slurry vessels, the total microbial population and DNT degradation population was enumerated. Enumeration on R2A and 2,4-DNT revealed cell concentration of $6.9\ 10^6$ and $1.7\ 10^7$ cells/ ml, respectively. No growth was observed on 2,6-DNT. In addition, no growth was observed when enumerations were performed in the presence of $HgCl_2$ (@ 50 mg/L. This suggests that biological activity in the control flasks was due to the unavailability of $HgCl_2$ (because of its binding to the soil or volatilization), rather than the outgrowth of a $HgCl_2$ resistant population. Subsequent to this finding, additional $HgCl_2$ dosing (up to 200 mg/L) was performed. Microbial activity was not halted by increasing the concentration of $HgCl_2$ to 200 mg/L on day 19. A progressive decline in NO_2^- concentration in the control flasks, as well as all other flasks, suggests the occurrence of nitrification.

The DNT profile in Figure 3, Panel A again indicates rapid desorption of the DNTs. Approximately 0.14 mM of desorbable 2,4-DNT was present in all flasks. The near stoichiometric release of nitrite observed in Panel B and Panel C coinciding with the disappearance of 2,4-DNT reveals that aerobic denitration

pathways were expressed. In Flasks 3,4 (indigenous microorganisms) and 5,6 (supplemented with JS922) an average of 0.22 mM of NO_2^--N is released. Approximately 80% of all the nitro groups available in 2,4-DNT were, therefore, released as of NO_2^--N. The remaining 20% of the of NO_2^--N are likely used for cell synthesis.

Figures 4 reveals that transformation of 2,6-DNT is limited. Phase II of the experiment examined the possibility of toxicity or inhibition effects of 2,6-DNT or 2,4,6-TNT. The XAD-4 Tenax beads added to flasks 1, 3, and 5 reduced the aqueous concentration of 2,6-DNT and TNT. Active inoculum of 2,6-DNT degrading JS 922 was added to the flasks 5 and 6. The addition of Tenax beads did not appear to enhance 2,6-DNT degradation. The addition of fresh inoculum (JS922) did not appear to enhance 2,6-degradation.

A slow decrease in TNT concentrations are observed in all three treatments (Figure 5). Slow desorption of TNT is observed in the bioaugmented control flasks relative to the abiotic and biotic controls. The same slow desorption is observed with 2,6-DNT in the bioaugmented control flasks.

Figure 6 shows the results of the TNT toxicity/inhibition experiment. The initial rate of 2,6-DNT removal was very similar in the control (0.015 mg L^{-1} hr^{-1}) and TNT supplemented (0.018 mg L^{-1} hr^{-1}) flasks. Thus, at a concentration of 25 mg/L TNT does not impact initial cell activity. Note, however, that only 10% of the initial 2,6-DNT was degraded in these experiments, and the effects of TNT on sustained 2,6-DNT transformation are not clear.

DISCUSSION

Soil B contained large amounts of 2,4-DNT, 2,6-DNT, and 2,4,6-TNT. The rapid desorption of the contaminants, suggests that the nitrotoluenes are not strongly bound to the soil organic matter or present deep in the soil micropores. Recently, it has been demonstrated that nitroaromatics can adsorb specifically and reversibly to clay surfaces (5). This adsorption occurs via an electron acceptor-donor complex formation between the electron deficient nitroaromatic nucleus and the electron dense oxygen ligands at the clay surface (5, 16). The dominant mineral in the examined soil was kaolinite (15). Therefore, the majority of the desorbable NTs were probably adsorbed to the clay surface rather than partitioned in the soil organic matter. Changes in ionic strength from 0.1 mM to 0.1 M have little effect, while the type of exchangeable cations can have a large effect on NT adsorption to clays (5). Thus, a difference in the ionic solution between the site geochemistry, which is unfortunately unknown, and the mineral salts solution used in these experiments may account for the rapid desorption observed in the lab *vs* their persistence in the field.

The indigenous microorganisms present in soil B were capable of complete mineralization of 2,4-DNT. Near stoichiometric release of nitrite accompanied the degradation of the desorbable 2,4-DNT. Because nitrite was released in near stoichiometric quantities, aerobic denitration pathways are inferred. This contrast with a recent report on aerobic 2,4-DNT transformations in soil microcosms by indigenous microorganisms where only 28% mineralization was observed, and 28% of the original compound was transformed to reduced intermediates (1). Because of the activity of the indigenous microbial community, addition of a known 2,4/2,6-DNT degrader (JS922) was not beneficial.

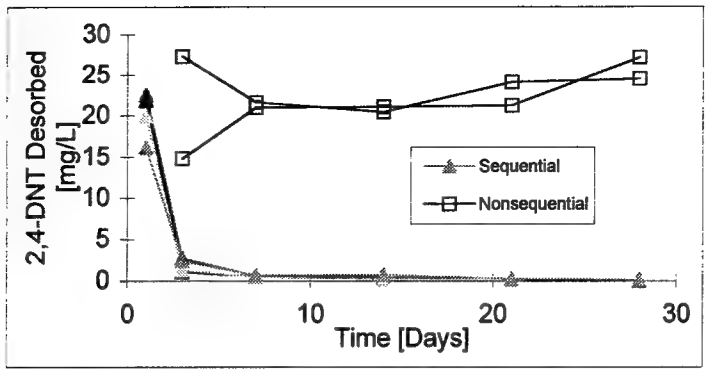
No mineralization of 2,6-DNT was found during these experiments. MPN enumeration on the soil, however, revealed a population with 2,6-DNT denitrating activity. A small fraction of 2,6-DNT was mineralized (8%) in the reported aerobic microcoms studies, but in that case TNT was absent and the initially applied DNT concentrations were not reported (1). Possibly, toxicity or inhibition exerted by the high aqueous concentrations of 2,6-DNT and/or 2,4,6-TNT impair 2,6-DNT biotransformation. After 25 days of incubation, MPN analysis was performed on a soil slurry supplemented with JS 922 and one that only contained the indigenous microflora. MPNs on R2A and 2,4-DNT for the indigenous and augmented slurry reactors were $3.5\ 10^7$ and $9.18\ 10^4$ and $4.62\ 10^7$ and $1.8\ 10^5$ cells/ml for respectively. Thus, a slightly higher active 2,4-DNT degrading population was present with JS 922 supplementation. However the measured 2,4-DNT degrading population was nearly 3 orders of magnitude below the initial JS 922 cell concentration. Enumeration on 2,6-DNT revealed no cell count for either condition. This suggests that either the added 2,6-DNT degraders have disappeared (because of toxicity of the environment), or are unrecoverable after 25 days on a medium containing 2,6-DNT as sole C source. Neither the addition of a sorptive matrix to decrease the aqueous concentration of 2,6-DNT and TNT in the soil slurry experiments (phase II), nor the addition of an induced culture (JS922) grown on 2,6-DNT enhanced 2,6-DNT degradation. MPN enumeration of the 2,4-DNT and 2,6-DNT degrading fraction was performed 10 days after the onset of phase II. 2,4-DNT and 2,6-DNT MPN were $4.9\ 10^5$ and $4.9\ 10^6$ cell/ml when XAD-4 was present, while those numbers were $3.6\ 10^3$ and $3.6\ 10^2$ in the absence of XAD-4. These results suggest that the reduction of aqueous phase TNT concentration increases the recoverability of DNT

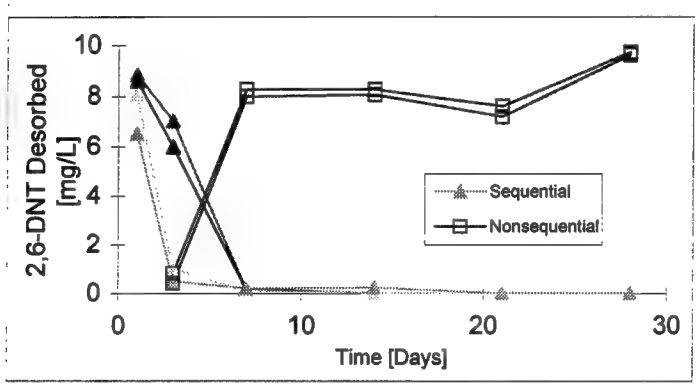
degrading organisms, and potentially the survivability of JS 922. However, 2,6-DNT mineralization was not observed (i.e. no NO_2^- production) even after the combined treatment of XAD-4 addition and a second inoculation, indicating that other factors impede activity of JS 922.

We postulate that the observed decrease in the aqueous phase concentration of 2,4,6-TNT (Figure 5) is not due to biotic transformations. No additional peaks reflective of aminodinitrotoluenes isomers were observed on the HPLC chromatographs suggesting that nitro reductive pathways were not significant. Furthermore, no organisms have so far been isolated that can aerobically mineralize TNT as sole carbon and nitrogen source. TNT loss by volatilization is also unlikely, as the Henry's law coefficient for TNT, 2,4-DNT and 2,6-DNT are 1.1 10^{-8}, 1.86 10^{-7}, and 4.86 10^{-7} atm m^3 mol^{-1} indicating that TNT is the least volatile of the three nitrotoluenes. The reasons for the TNT disappearance remain, therefore, elusive.

REFERENCES

1. **Bradley, P. M., F. H. Chappelle, J. E. Landmeyer and J. G. Schumacher.** 1997. Potential for intrinsic bioremediation of a DNT-contaminated aquifer. Groundwater **35:**12-17.
2. **Carmichael, L. M., R. F. Christman and F. K. Pfaender.** 1997. Desorption and mineralization kinetics of phenathrene and chrysene in soils. Environ. Sci. Technol. **29:**126-132.
3. **Carroll, K. M., M. R. Harkness, A. A. Bracco and R. R. Balcarel.** 1994. Application of a permeant/polymer diffusion model to the desorption of polychlorinated biphenyls from Hudson river sediments. Environ. Sci. Technol. **28:**253-258.
4. **Eaton, A. D., L. S. Clesceri and A. E. Greenberg.** 1995. Standard Methods for the Examination of Water and Wastewater. American Puclic Health Association,Washington, DC
5. **Haderlein, S. B., K. W. Weissmahr and R. P. Schwarzenbach.** 1996. Specific adsorption of nitroaromatic explosives and pesticides to clay minerals. Environ. Sci. Technol. **30:**612-622.
6. **Haigler, B. E. and J. C. Spain.** 1996. Degradation of nitroaromatic compounds by microbes. SIM News **46:**59-68.
7. **Jenkins, T. F., M. E. Walsh, P. W. Schumacher, P. H. Miyares, C. F. Bauer and C. J. Grant.** 1989. Liquid chromatographic method for determination of extractable nitroaromatic and nitramine residues in soil. J. Assoc. Off. Anal. Chem. **72:**890-899.
8. **Klee, A. J.** 1993. A computer program for the determination of most probable numbers and its confidence limits. J. Microb. Meth. **18:**91-98.
9. **Nishino, S. F.** 1986. Direct acridine orange counting of bacteria preserved with acidified lugol iodine. Appl. Environ. Microbiol. **52:**602-604.
10. **Ramaswami, A. and R. G. Luthy.** 1997. Measuring and modeling physicochemical limitations to bioavailability and biodegradation, p. 721-729. *In* C. J. Hurst, G. R. Knudsen, M. J. McInerney, L. D. Stetzenbach and M. V. Walter (ed.), Manual of Environmental Microbiology, ASM Press, Washington, DC.
11. **Reasoner, D. J. and G. E. E.** 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. **49:**1-7.
12. **Spain, J. C.** 1995. Biodegradation of nitroaromatic compounds. Ann. Rev. Microbiol. **49:**523-555.
13. **Spanggord, R. J., J. C. Spain, S. F. Nishino and K. E. Mortelmans.** 1991. Biodegradation of 2,4-dinitrotoluene by a *Pseudomonas* sp. Appl. Environ. Microbiol. **57:**3200-3205.
14. **Stanier, R. Y., N. J. Palleroni and M. Doudoroff.** 1966. The aerobic pseudomonads: A taxonomic study. J. Gen. Microbiol. **43:**159-171.
15. **TRW.** 1996. Site characterization report for the Volunteer Army Ammunition Plant National Environmental Technology Test Site. Site characterization report no. SFIM-AEC-ET-CR-96154. US Army Environmental Center.
16. **Weissmahr, K. W., S. B. Haderlein, R. P. Schwarzenbach, R. Hany and R. Nüesch.** 1997. *In situ* spectroscopic investigations of adsorption mechanisms of nitroaromatic compounds at clay minerals. Environ. Sci. Technol. **31:**240-247.





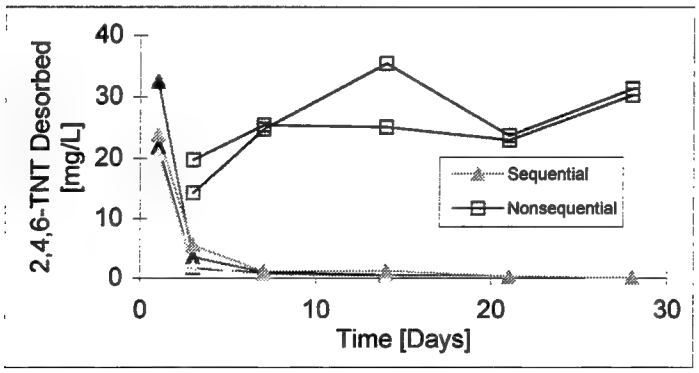


Figure 1 - Desorption of 2,4-DNT (Top), 2,6-DNT (Center), and 2,4,6-TNT (Bottom) from Soil B using XAD-4 Tenax Beads.

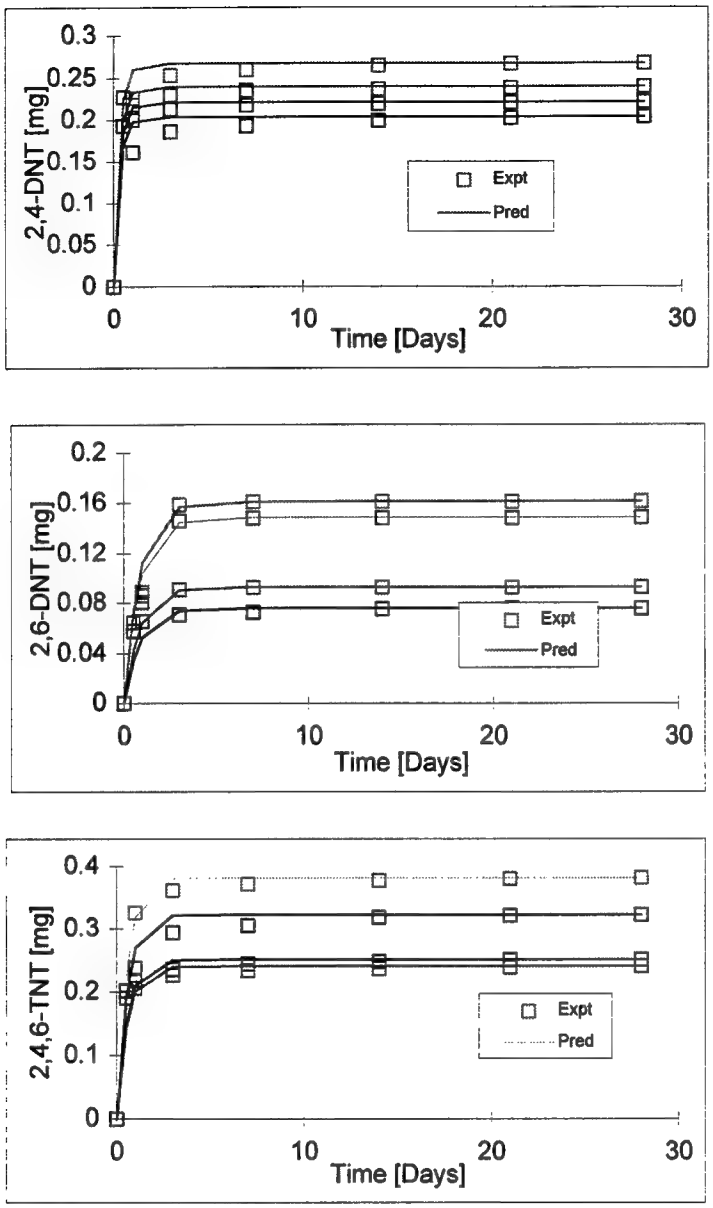
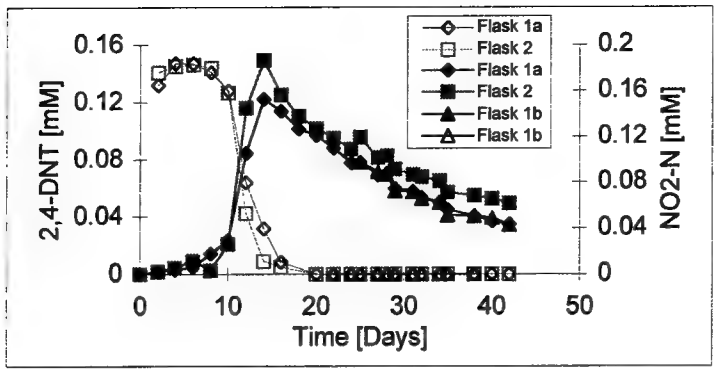
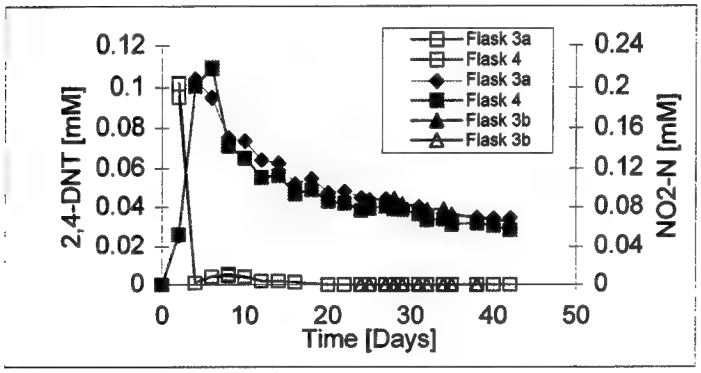


Figure 2 - Cumulative desorption of 2,4-DNT (Top), 2,6-DNT (Center), and 2,4,6-TNT (Bottom) from Soil B using XAD-4 Tenax Beads. Data fit used a first order desorption model.





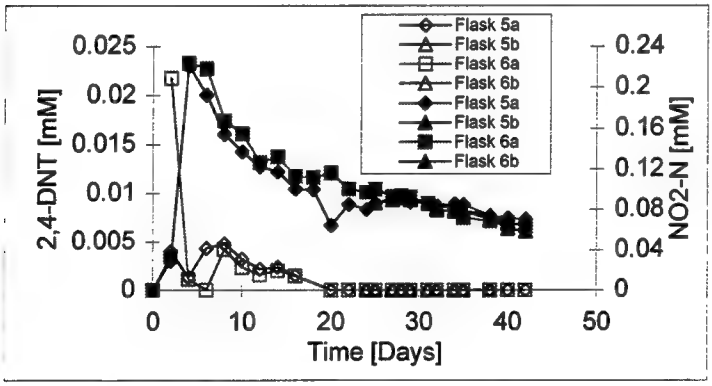


Figure 3 - 2,4-DNT (open symbols) and nitrite (closed symbols) profiles in abiotic control (Top), biotic control (Center), and JS922 supplemented tests (Bottom).

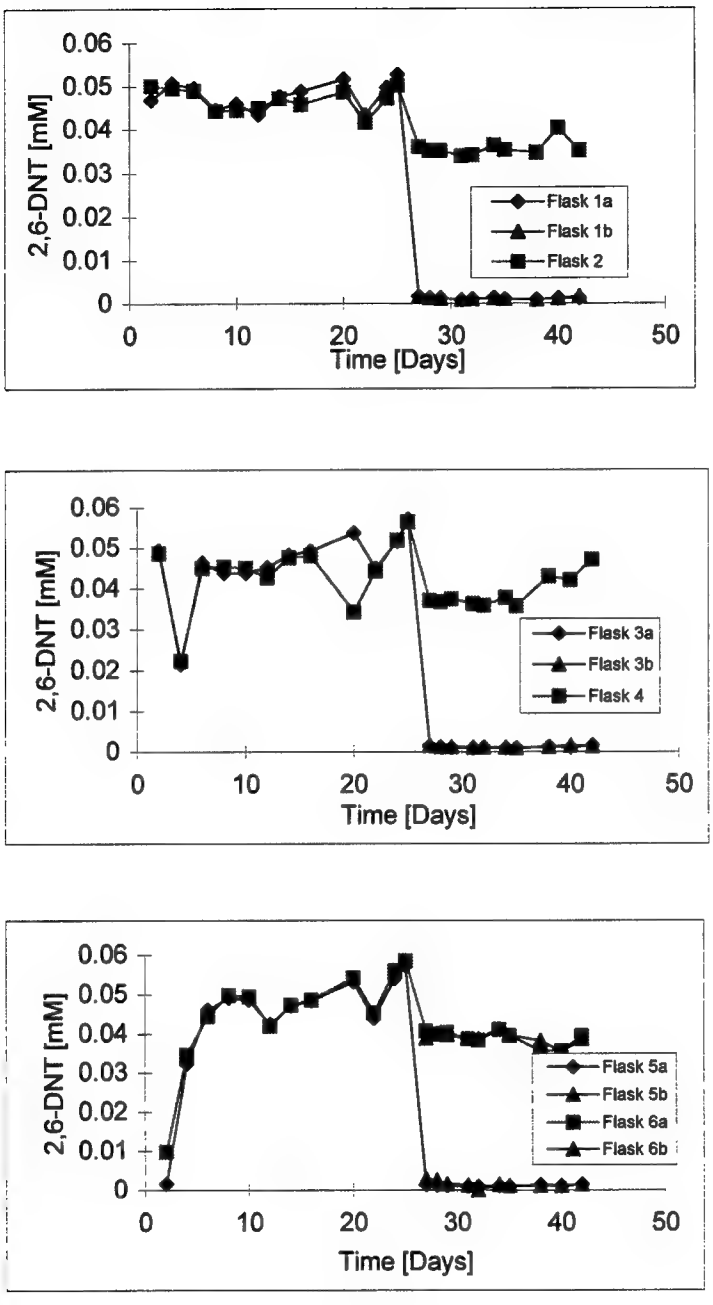
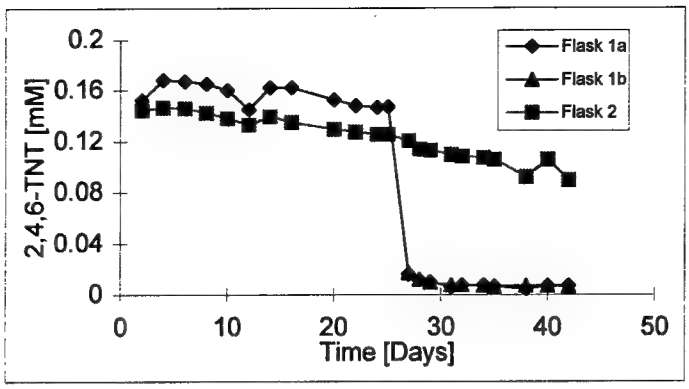
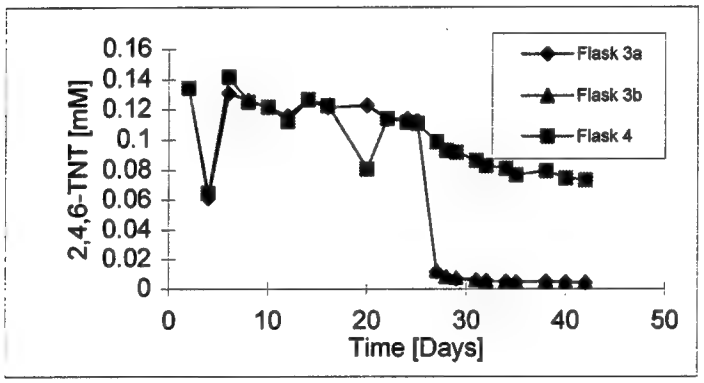


Figure 4 - 2,6-DNT profiles in abiotic control (Top), biotic control (Center), and JS922 supplemented tests (Bottom).





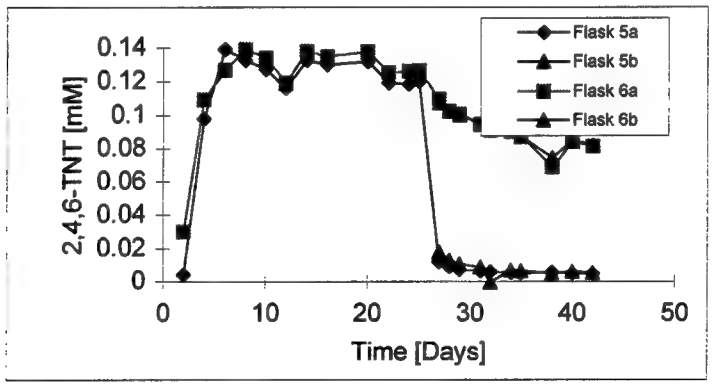


Figure 5 - TNT profiles in abiotic control (Top), biotic control (Center), and JS922 supplemented tests (Bottom).

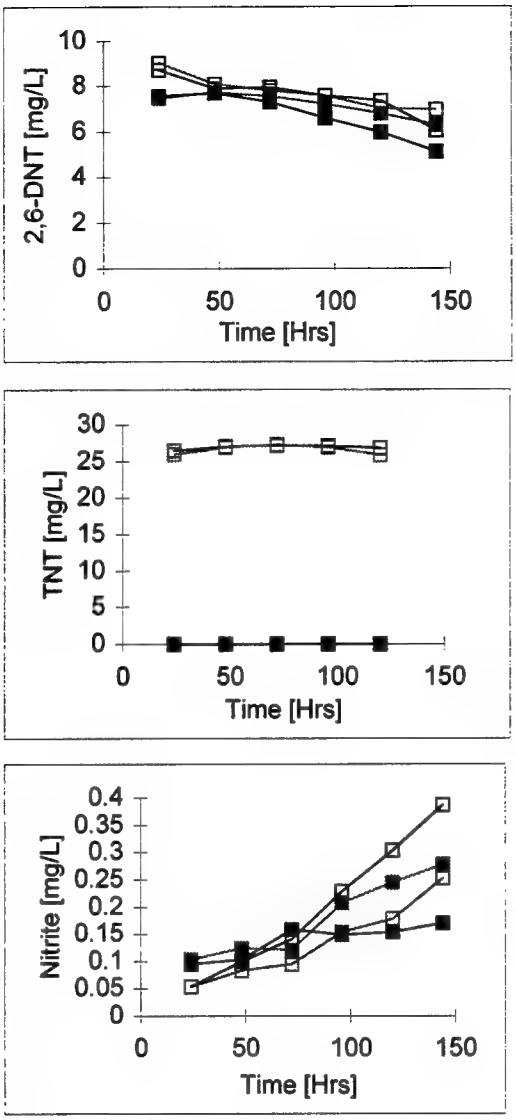


Figure 6 - 2,6-DNT (Top), TNT (Center), and nitrite-N (Bottom) profiles during TNT toxicity/inhibition experiments. Open and filled symbols refer to TNT supplemented and non-supplemented flasks, respectively.

IMPLEMENTATION OF FREFLEX/MERLIN TELEOPERATION

Mark A. Murphy
Graduate Student
Department of Mechanical Engineering

Ohio University
257 Stocker Center
Athens, OH 45701-2979

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

IMPLEMENTATION OF FREFLEX/MERLIN TELEOPERATION

Mark A. Murphy
Graduate Student
Department of Mechanical Engineering
Ohio University

Abstract

This report summarizes the author's work in the AFOSR 1997 Graduate Student Research Program, based in the Human Sensory Feedback Laboratory of Armstrong Laboratory, located at Wright-Patterson AFB. A powerful and general control architecture is being implemented for real-time, sensor-based, rate-based, shared control of a *Merlin* industrial robot with the unique *Freflex* force-reflecting exoskeleton. This work is the first time the *Freflex* exoskeleton has been used for teleoperation of the *Merlin* slave manipulator. A description, including operating procedures, of the teleoperation resources of the Laboratory is included within the report. Additional accomplishments are the novel Naturally-Transitioning Rate-to-Force Controller (NTRFC) and a *Matlab* simulation of *Freflex/Merlin* teleoperation under joint and Cartesian pose and rate control.

IMPLEMENTATION OF FREFLEX/MERLIN TELEOPERATION

Mark A. Murphy

1 Introduction

Teleoperation of remote manipulators is greatly enhanced by using a force-reflecting input device. This force/moment haptic feedback increases the sense of telepresence (where the user feels part of the remote or virtual environment) by enabling the operator to feel through the force-reflecting master the forces and moments exerted by the slave manipulator on the environment. The Human Sensory Feedback (*HSF*) Laboratory of Armstrong Laboratory located at Wright-Patterson AFB has a world-class capability for experimentation in force-reflecting teleoperation for Air Force and NASA applications: The unique *Freflex* force reflecting exoskeleton master (Odetics, 1992), a *Merlin* industrial manipulator slave (American Robot Corporation, 1985), and a *Chimera* 3.2 real-time operating system for reconfigurable sensor-based control (Ingimarson, 1995) loaded on a Sun SPARCstation. The purpose of this report is to document the author's summer 1997 work at the *HSF* Lab in *Freflex/Merlin* force-reflecting teleoperation implementation, sponsored by the AFOSR Summer Graduate Student Research Program.

The telerobotic control theory applied in the current report was developed by Dr. Robert Williams II, an Ohio University Assistant Professor who advised the author during the summer 1997 work in the HSF Lab. His companion report (Williams, 1997b) focuses on this telerobotic control architecture. For a more complete report of the summer's work, please see Williams and Murphy (1997a). The author accomplished two basic results in summer 1997 work: 1) Implementation of *Freflex/Merlin* teleoperation (currently in progress). Simulation of *Freflex/Merlin* teleoperation and assistance of development of control architecture for general telerobotic systems including force-reflecting hand controllers (FRHCs). 2) Controls design, modeling, simulation, and draft journal article (Williams and Murphy, 1997b) for the novel Naturally-Transitioning Rate-to-Force Controller (NTRFC),

which is part of 1). This report summarizes the implementation since the remaining accomplishments are detailed in Williams and Murphy (1997b) and the companion report of Dr. Williams II (1997b).

2 System Description

With the possession of the *Freflex* (*F*orce-*refl*ecting *ex*oskeleton, Fig. 1, a unique device from an Odetics, Inc. SBIR, 1991), *Merlin* (*M*odular *e*xpandable *r*obot *lin*e, Fig. 2, from the American Robot Corporation, 1985), and the *Chimera* 3.2 real-time operating system, the HSF Laboratory of Armstrong Laboratory located at Wright-Patterson AFB has a world-class capability for experimentation in force-reflecting teleoperation for Air Force and NASA applications.

Figure 1. Freflex Force-Reflecting Exoskeleton

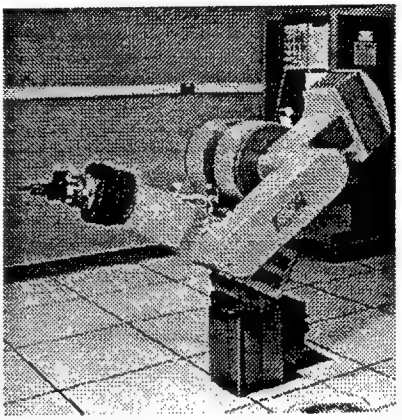


Figure 2. Merlin Slave Manipulator

2.1 HSF Laboratory Hardware

The Human Sensory Feedback (HSF) Laboratory of Armstrong Laboratory at Wright-Patterson AFB has the following hardware devices for teleoperation studies. Telerobotic control (see Williams, 1997b for the theoretical architecture) was implemented on these devices in simulation and hardware implementation is currently progressing.

2.1.1 *Merlin.* The *Merlin* system includes a six-dof arm, controller, peripherals (teach pendant, CRT, and operator's panel), and AR-BASIC software.

The left-handed *Merlin 6500* robot arm located in the HSF Laboratory is balanced for a payload of 50 lb. The six joints of the *Merlin* are driven by stepper motors. Each time the motors receive a current pulse from the motor drivers, the motors step by 1/25,000 of a motor revolution. Encoders are mounted on the back of each motor

to determine the position of each motor. These encoders read 4000 ticks per revolution and the controller counts the number of motor revolutions to calculate precise joint positions.

The adjustable mechanical stops, the electromagnetic brakes which prevent the arm from moving when the robot power is off, and the transmissions for the waist, shoulder, and elbow joints are located in the body of the robot. Electrical stops are located in the body for the waist and shoulder joints, while the limit switch for the elbow is located in the upper-arm. The transmissions for the wrist axes consist of three concentrical tubes that pass through the forearm and transmit torque from the motors to the wrist joints. The range of motion due to the mechanical stops and the gear ratios for each joint are shown in Table I.

Table I. Merlin Range of Motion due to Mechanical Stops and Gear ratios for each Joint

Axis	Range of Motion	Gear Ratios
Waist	294 degrees	48 : 1
Shoulder	292 degrees	48 : 1
Elbow	292 degrees	48 : 1
Wrist Rotate	+/- continuous	24 : 1
Wrist flex	+/- 90 degrees	20 : 1
Hand	+/- continuous	24 : 1

The *Merlin* controller consists of two PC's, the Master CPU and the Servo CPU (see American Robot Corp., May, 1995). The Master CPU is an MSDOS based PC which runs a program that contains the AR-Basic interpreter as well as the motion libraries, ARBASIC.EXE. The Servo CPU is actually a single board ISA processor with RAM, and VGA that runs SERVO.COM which is responsible for the control of the motors. Both of these CPUs are located in the PC within the American Robot Control Cabinet. The High Speed Host Interface (see American Robot Corp., 1997) is implemented to allow the user to communicate directly with the Servo CPU via an RS-232 interface.

The three peripherals that are most often used in the HSF Laboratory are the teach pendant, CRT, and the operator's panel. The teach pendant (American Robot Corp., July, 1995), which is connected to the front of the

control cabinet, is used for manual control of the *Merlin*. The operator's panel primary functions are emergency stopping, electromagnetic brake override, robot power activation, and controller power activation. The main purpose of the CRT is to display the AR-Basic window.

AR-Basic is a version of the programming language BASIC developed by American Robot Corp. to create programs to command the *Merlin* (American Robot Corp., 1996).

2.1.2 *Freflex* The master is the seven-dof, seven-*R Freflex* (Odetics,1991). This unique exoskeleton anthropomorphically maps the seven-dof of the human right arm and with force-reflection provides a natural means to teleoperate a remote manipulator. The human is not attached to the exoskeleton which allows the user to feel less constrained. A forearm push plate can be enabled to maintain a constant force on the user's arm to better follow the motions of the human.

The *Freflex* is driven by seven brushless p.m. servomotors which are controlled by BDS4 series brushless motor controllers (Industrial Drives, Division of Kollmorgen Corporation). These motors can provide high continuous torque and low armature friction and inertia. Bayside gearheads with reductions ranging from 5:1 to 15:1 are mounted on each motor (Odetics, 1991). The motors are mounted on an external base minimizing the size, mass, and inertial properties of the *Freflex* exoskeleton.

The exoskeleton has a cable transmission consisting of 19 shafts, 102 pulleys, 92 bearings, and a gear set at the elbow. The cables, acting as antagonistic tendons, are routed from the motors to the joints of the exoskeleton via the pulleys mounted at the base. This cable transmission causes the motion of the *Freflex* joints to be coupled (Huang, 1993).

The seven joint angles of the *Freflex* are measured by external potentiometers that are designed into the joint such that the wiper is mounted on the support link and the film is mounted on the driven link. According to Odetics, this arrangement provides better packaging and positive joint angle measurements and requires fewer calibrations. However, research in the HSF lab indicates that these pots are the greatest liability for dependability on the *Freflex*. These pots are unique and have failed in the past. Currently the pot on joint three does not work as designed. A wire has been temporarily added from the brush to the red wire leaving the pot to bypass a discontinuity in the original red wire exiting the pot. To calibrate the pots see section 4.1.

The *Freflex* VME chassis contains four VME-based processors and I/0 boards mounted in a 21 slot Chassis. The VMIC 4100 board outputs voltages to the *Freflex* motor controllers. The VMIC 2510B board provides discrete input and output channels for the exoskeleton operator interface. The Ironics IV-3230 board labeled "crusher" is used for force-reflection processing while the Ironics IV-3230 board "control" is the Master Real-Time Processing Unit. The chassis also contains a JR3 board that processes information from the JR3 force/torque sensor mounted at the wrist of the *Freflex*, a Data Translation DT1401 card that reads the pots, and a Bit 3 card that links the enet and the Sun SPARCstation.

2.2 HSF Laboratory Software

Chimera 3.2 (a real-time operating system for reconfigurable sensor-based control systems developed by the Advanced Manipulators Laboratory, The Robotics Institute, and the Department of Electrical and Computer Engineering at Carnegie Mellon University), loaded on a Sun SPARCstation, acts as the interface between the *Freflex* and the *Merlin*.

2.2.1 *Chimera* 3.2 This section of the report will try to summarize the necessary material concerning the Chimera 3.2 real-time operating system needed to operate the *Merlin* manipulator via the *Freflex* exoskeleton. Note that an understanding of UNIX commands is essential since Chimera 3.2 is accessed from a Sun SPARCstation.

2.2.1.1 Subdirectory Structure. The base directory for controlling the *Merlin* with the *Freflex* is named freflex_merlin and it is located at /usr/chimera/chim_3.2/freflex_merlin. Chimera 3.2 defines a subdirectory Hierarchy that all files under the base directory must follow (Ingimarson, 1995). Fig. 3 shows this hierarchy under freflex_merlin.

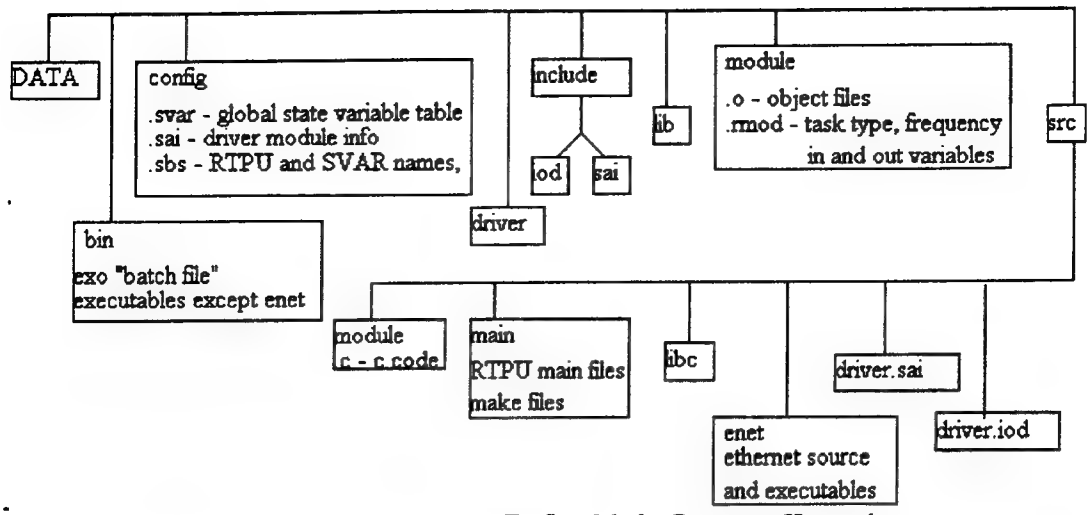


Figure 3. Freflex_Merlin Directory Hierarchy

2.2.1.2 Editing, Compiling, and linking Files. There are two file editors available on the Sun/UNIX system, the vi editor and nedit. To compile a single module use the makefile in the subdirectory in which the file is located, adding the filename if not already on the list of object files to be created within the makefile. The command "make filename.o" should be used to compile a single module. If .o is not applied to the end of the filename an error will result. To compile and link all modules for RTPU "crusher" type "make -f freflex_merlin_rtpu.mak all". Change the directory to /usr/chimera/chim_3.2/freflex_merlin/src/main, remove "freflex_merlin.o", and type "make all" when compiling and linking all modules for RTPU "control". A Chimera 3.2 library with helpful functions, including a matrix math library which was used extensively in the freflex_merlin development, is available. These libraries are detailed in the Chimera 3.2 manual.

2.2.1.3 Reconfigurable Subsystems. Each periodic control module is composed of several components as required by Chimera 3.2. These components are the C subroutines: *init, on, cycle, off, kill, clear, set,* and *get* (Ingimarson, 1995). Existing .c files in the /usr/chimera/chim_3.2/freflex_merlin/src/module directory are good examples of the subsystem (SBS) configuration. When operating Chimera 3.2 there is a SBS interactive command interpreter that is useful. It accepts the commands: *display, kill, module, off, on, quit, set, spawn*, and *status* (Ingimarson, 1995). Details of these commands can be found in the Chimera 3.2 manual or by typing "help" when operating Chimera 3.2.

2.2.1.4 Global State Variable Table. Chimera 3.2 uses a global state variable table (SVAR), which is stored in a configuration file, to communicate between different modules when operating in a reconfigurable multiprocessor environment(Ingimarson, 1995). Each task creates and modifies a local copy of the SVAR and periodically returns these global variables, which are used by different tasks, back to the SVAR.

3 Hardware Control Under Chimera 3.2

The control flow diagrams for the *Merlin* using the *Freflex* and for the *Freflex* independently are presented in this section.

3.1 Merlin Control

The *Merlin* will possess the ability to be controlled by three general modes (Williams, 1997b), joint, resolved-rate, or pose (position and orientation). Implementation is almost complete for resolved-rate control while the installation of pose and joint control is progressing. Fig. 4 shows a flow diagram for *Merlin* operation via the *Freflex*.

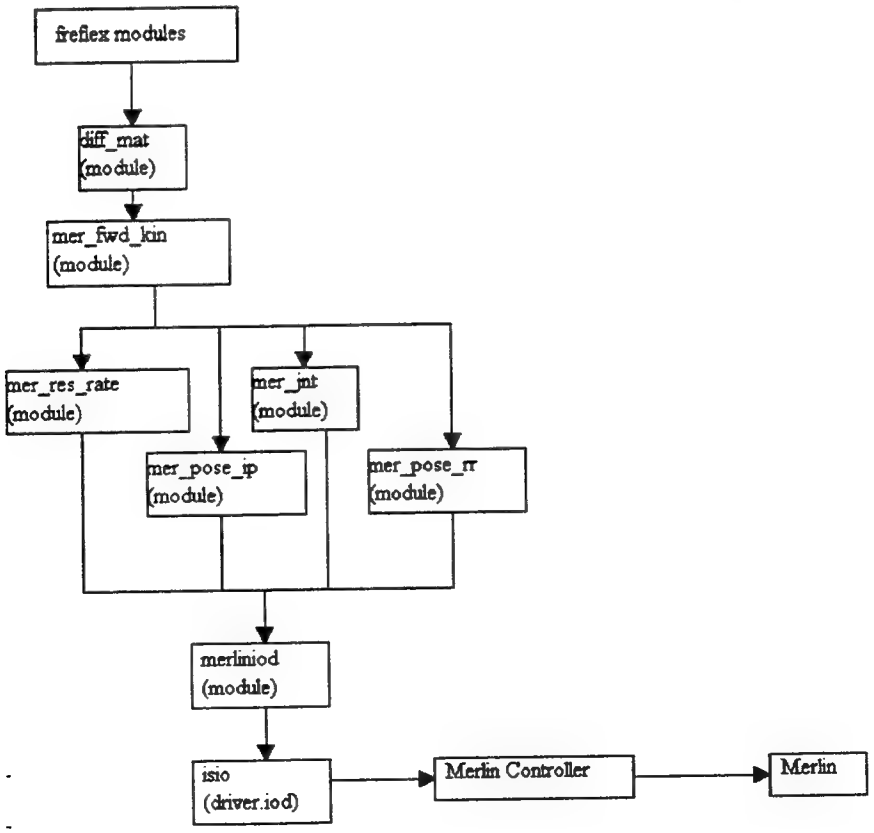


Figure 4. Freflex Control of Merlin under Chimera 3.2

3.2 Freflex Control

A flow diagram for the *Freflex* is shown in Fig. 5.

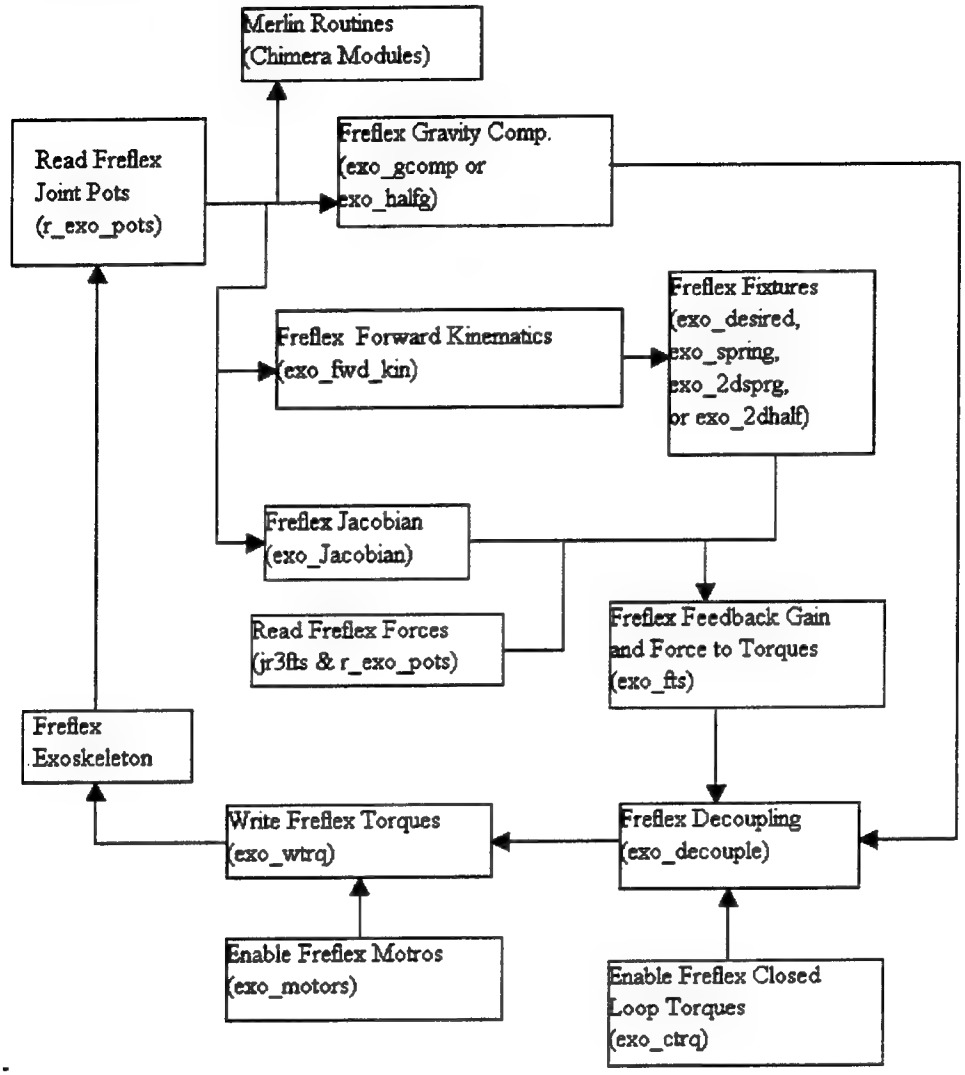


Figure 5. Freflex Control under Chimera 3.2

4 Operating Procedures

Various operating procedures for the *Freflex* and *Merlin* are presented in this section.

4.1 Freflex

4.1.1 Hardware Power Up The following steps illustrate the power up procedure for the *Freflex* exoskeleton.

1. Turn on the circuit breaker in the power panel labeled:

   PPL

   120-208V 3@ 4W

   FM PDP-A

   The circuit breaker is labeled *Freflex* at positions 8, 10, and 12.

2. Power up the VME chassis at the rear of the cabinet.

   Note: Failure to connect the d/a cable to VMIVME board 4100 has resulted in *Freflex* damage in the past.

3. Power up the *Freflex* interface chassis at the front of the unit which is located above the VME chassis.

   Note: The emergency stop button should be out and if the fault led is lit press the reset button.

4. Power up the *Freflex* base unit by activating the circuit breaker that is located to the right of the power cord on the base.

   Note: Press the yellow reset switch/lamp on the base if lit for operation. For activation, the foot and the grip switches must be depressed and the software controlled safety relay must be running. The signal connectors below the *Freflex* grip should be checked before operation since these are easily disconnected and prevent the activation of the exoskeleton.

5. Undo steps 4 through 1 to shutdown the system.

4.1.2 Software Start Up The software for operating the *Freflex* in Chimera 3.2 can now be started using the following procedure:

1. Login under the Chimera account on the SUN SPARCstation.
2. From an xterm window: Change to the freflex_merlin binary directory by typing "cd chim_3.2/freflex_merlin/bin" (home directory is /usr/chimera). Start Chimera 3.2 by typing "chim" and execute the batch file to download both processors by typing "<exo" at the chim prompt.
3. Start the enet socket window for data display by typing "cd/usr/chimera/chim_3.2/freflex_merlin/src/enet" from another xterm window. Activate the data display program by entering "enet_disp" in this window.

Note: The enet display needs to be started before RTPU "control" is activated since "control" spawns the sunenet routine which tries to connect to "enet_disp". Start "enet_data" instead of "enet_disp" to collect and save data to the SUN disk. This step must be completed before the "sundata" routine is started after loading the freflex_merlin program in step 2 since the sundata routine tries to connect to "enet_data". Finally start the data transfer routine, "sundata", by typing "on sundata" at the "chim" prompt.

4. Start execution on "control" by typing "go control" at the "chim" prompt. Type "status" to see which programs are running and type "on *filename*" and "off *filename*" to turn programs on and off respectively.
5. Type "quit" at the "chim" prompt to terminate the programs and disable the *Freflex*. Typing "quit" again will terminate Chimera 3.2.

4.1.3 Troubleshooting The qs server may remain running if the system ends abnormally. The program "chimclean" can be used to clean-up the system and terminate the qs server by typing "chimclean" at the operating system prompt.

If the system has been rebooted, there is a message when Chimera 3.2 is started that says "sxm drivers are not installed", or if the system locks up while trying to start Chimera 3.2, then the sxm drivers need to be installed using the following procedure:

1. Change to the Sun4c driver directory in an xterm window by typing "cd /usr/chimera/chim_3.2/sun4c/src/driver".
2. Login as the superuser.
3. Type "sxm.INSTALL" at the superuser prompt "%" to execute the batch file.
4. If no error messages are displayed and the display indicates that the sxm driver has been loaded, then type "exit" to logout as superuser.
5. Restart the Chimera 3.2 software.

If the system operates but the data output is drifting or not acting as expected, the pots should be checked for proper calibration and operation. While the system is running, place the *Freflex* in a fixed position. Select the

pots enet display screen by typing "2" on the enet display. If any of the angles are drifting, the pot is not working properly. Use a voltmeter on the pot to identify the problem.

To check the pot calibration, rotate each joint until the calibration marks on the joint line-up. Joints one, three, and five should read zero degrees, joints two and six should read 90 degrees, and joint four should read -90 degrees when at their respective calibration positions. If a pot needs recalibrated, record the hexadecimal A/D number displayed for the corresponding pot on the enet "forces" display screen by typing "4" on the enet display. Then open the file "jointpot.sai" in the directory /usr/chimera/chim_3.2/freflex_merlin/config and change the "cal_cnts" number for the corresponding pot to the recorded hexadecimal number.

4.2 Merlin

4.2.1 Powering up the *Merlin* The procedure for powering up the *Merlin* is listed below:

1. Turn on the circuit breaker, switch 21, in the power panel on the east wall labeled:

   PPL

   120-208V 3@ 4W

   FM PDP-A

2. Ensure that the CRT and keyboard are connected to the PC inside the *Merlin* control cabinet and that the CRT is powered up.
3. Turn on the circuit breaker at the bottom right-hand side on the front of the *Merlin* control cabinet.
4. Pull out the emergency power button on the operator's panel on the front of the cabinet.
5. Key at bottom of operator's panel needs to be set to "run".
6. Depress the POWER ON button on the operator's panel.
7. Depress the MOTOR POWER button on the operator's panel.

   Note: The PC controller should now boot up in MS-DOS within a few seconds.

8. At the C:> prompt, type "I" to load the second processor and to start the server. Type "A" to start AR-Basic.

   Note: If Ctrl-Alt-Del is used to reset the computer, the second processor does not need to be reloaded.

9. Type "robot on" to set the robot to online.

10. Type "calibrate init" at the command prompt within AR-Basic. While the light fence is activated and the foot pedal is depressed, use the teach pendant to move the waist to the middle of the desired operating range and position the arm and wrist straight-out. To move the *Merlin* arm using the pendant, the button "jnt" must be pressed and then the button "pos" or "ori" must be held while using the three control bars on the right side of the pendant to actuate the desired joint. Press the "F1" function key on the keyboard while maintaining the light fence active and the foot pedal depressed. When the command prompt reappears, the calibration is complete.
11. To power off, type 'exit' in AR-Basic and undo steps 3 and then 1.

4.2.2 Troubleshooting If the wrist joints are emitting "grinding" sounds while moving and not moving smoothly, check that the robot is receiving proper plant air pressure.

If the robot encounters a limit switch during operation, it can go off-line. To alleviate this problem turn the key on the operator's panel to "BR OV" to override the brake and then manually rotate the robot into its workspace. Return the key to "RUN" to operate the *Merlin*.

4.3 Freflex Commanding Merlin

The procedure for controlling the *Merlin* manipulator via the *Freflex* is listed below:

1. Follow steps 1 through 10 for powering up the *Merlin*.
2. Exit from AR-Basic by typing "exit" and type "hshicom" at the C:>AR-Basic> prompt.
3. Follow steps 1 through 4 for powering up the *Freflex* hardware.
4. Follow steps 1 through 4 for starting the *Freflex* software. Depress the *Merlin* foot pedal and turn on the driver by typing "on merliniod".
5. To shutdown the system via the serial link from the controlling processor, type "exit". If HSHI locks up, Ctrl-Alt-Del will reboot MS-DOS. Reverse steps 3 and 1 for powering up the *Merlin*. Follow step 5 from section 4.1.2 and then step 5 from section 4.1.1 to shutdown the *Freflex*.

5 Continuing Implementation

This section documents the current status and future goals for implementation. Gravity compensation is operational and the addition of force reflection is planned for future implementation in both the pose and rate control modes.

5.1 Rate Control

The resolved-rate control algorithm (Williams, 1997b) with gravity compensation is currently implemented for *Freflex* control of the *Merlin*. Future additions to this control method include a force controller for implementation of the NTRFC (Williams and Murphy, 1997b) system and constant-force return-to-center (CFRTC) and walls (Williams, 1997b).

5.2 Pose Control

Two types of pose control are being implemented for *Freflex* control of the *Merlin*.

5.2.1 Using Resolved-Rate Algorithm. One method of pose control is a method using the resolved-rate algorithm (Williams, 1997b). This method uses a closed position loop around the resolved rate system. This method will be enhanced by an index button, which is necessary since the *Merlin's* workspace is much larger than that of the *Freflex*, located on the grip of the *Freflex*. The button will allow the user to maneuver throughout the workspace of the *Merlin* by resetting the corresponding position of the *Freflex* to its current position at the time of index actuation.

5.2.2 Using Inverse-Position The second method of pose control being implemented is the traditional inverse-position control. This pose control method will possess the same index button as the pose control option listed in section 5.2.1.

5.3 Joint Control

Joint control is beneficial to the teleoperator controlling a kinematically dissimilar manipulator. This control method allows the operator to move desired links to avoid obstacles. Two types of joint control are being implemented to control the *Merlin* via the HSHI. It is not intuitive to use the *Freflex* to perform joint control of the kinematically dissimilar *Merlin*. Most likely the input joint commands to the *Merlin* will originate from the keyboard of the SUN SPARCstation.

5.3.1 Joint Position In this mode the user can command the position of a specific joint. This mode will rarely be used in teleoperation except for obstacle avoidance.

5.3.2 Joint Rate This mode allows the user to command an angular rate to a specified joint. This mode will also have limited teleoperation applications.

6 Conclusions

The Human Sensory Feedback Laboratory of Armstrong Laboratory, located at Wright-Patterson AFB, has great resources for teleoperation research. A powerful and general control architecture is being implemented for real-time, sensor-based, rate-based, shared control of a *Merlin* industrial robot with the unique *Freflex* force-reflecting exoskeleton. This work is the first time the *Freflex* exoskeleton has been used for teleoperation of the *Merlin* slave manipulator. The force reflection capability of the *Freflex* allows increased telepresence in remote operations. Additional accomplishments are the novel Naturally-Transitioning Rate-to-Force Controller (NTRFC) and a *Matlab* simulation of *Freflex/Merlin* teleoperation under joint and Cartesian pose and rate control.

Acknowledgments

The author gratefully acknowledges support for this research from the Air Force Office of Scientific Research and Armstrong Laboratory in the Summer Graduate Student Research Program. Many thanks also to Captain Debra North for support and facilities in the Human Sensory Feedback Lab at Wright-Patterson AFB. The following researchers helped make the summer work more productive and enjoyable: Lieutenant Kurtis Johnson, Dr. Dan Repperger, Jim Berlin, and Lieutenant Mike Krier. The author thanks Dr. Robert Williams II for the unlimited guidance provided through the teleoperation control theory and implementation.

References

American Robot Corporation, 1985, "*Service manual for the System II Merlin Intelligent Robot*".

American Robot Corporation, May, 1995, "*AARM Motion™ Retrofit for the Merlin® Industrial Robot MUCII 64" Controller, User and Service Manual*", Issue 1.1.

American Robot Corporation, July, 1995, "*AARM Motion™ Industrial Robot Controller, Teach Pendant User Manual*", Issue 1.2.

American Robot Corporation, 1996, "AR-Basic™ *V6.0 for the AARM Motion™ Industrial Robot Controller, Programming Manual*", Issue 1.7.

American Robot Corporation, 1997, "*HSHI High Speed Host Interface, Reference Manual*", Issue 2.0.
M.D. Bryfogle, 1990, "*Force Reflection Algorithms for Exoskeleton Controllers*", AAMRL-TR-90-090.
J.J. Craig, 1989, Introduction to Robotics: Mechanics and Control, Addison Wesley Pub. Co., Reading, MA.
C.J. Hasser, 1995, "*Tactile Feedback for a Force-Reflecting Haptic Display*", AL/CF-SR-1996-0134.
N. Hogan, 1985, "*Impedance Control: An Approach to Manipulation*", J. Dyn. Sys., Meas., and Cont.
M.Z. Huang, 1993, "*Efficient Coordination of an Anthropomorphic Telemanipulation System*", AL/CF-TR-1995-0120.
Industrial Drives, Division of Kollmorgen Corporation, *"Installation and Service Manual BDS4 Series Brushless Motor Controllers"*, M-89032, Issue 3.
Ingimarson, D., Stewart, D.B., Khosla, P.K., 1995, Chimera 3.2 The Real-Time Operating System for Reconfigurable Sensor-Based Control Systems, The Robotics Institute, Dept. of Electrical and Computer Engineering at Carnegie Mellon University, Pittsburgh, Pa.
Kane, T.R., Likins, P.W., and Levinson, D.A., 1983, Spacecraft Dynamics, McGraw-Hill.
Odetics Inc., 1992, "*Exoskeleton Master Arm, Wrist, and End-Effector Controller with Force-Reflecting Telepresence*", Phase II Final Report, Contract No. F33615-89-C-0587.
R.P. Paul, 1981, Robot Manipulators, MIT Press, Cambridge, MA.
M. Raibert and J.J. Craig, 1981, "*Hybrid Position/Force Control of Manipulators*", ASME Journal of Dynamic Systems, Measurement, and Control.
D.W. Repperger, E.L. Scarborough, and T.L. Chelette, 1991, "*Construction of a Dual Axis Force Reflection Stick and Test Station*", AL-TR-1992-0041.
D.W. Repperger, 1991, "*Active Force Reflection Devices in Teleoperation*", IEEE Control Systems Magazine, 52-56, January.
D.W. Repperger, C.A. Phillips, and T.L. Chelette, 1995, "*A Study on Spatially Induced "Virtual Force" with an Information Theoretic Investigation of Human Performance*", IEEE Transactions on Systems, Man, and Cybernetics, 25(10):1392-1404.
D.W. Repperger, 1995, "*Biodynamic and Spasticity Reduction in Joystick Control via Force Reflection*", AL/CF-TR-1995-0152.
L.B. Rosenberg, 1992, "*The Use of Virtual Fixtures as Perceptual Overlays to Enhance Operator Performance in Remote Environments*", AL/CF-TR-1994-0089.
L.B. Rosenberg, 1993, "*The Use of Virtual Fixtures to Enhance Operator Performance in Time Delayed Teleoperation*", AL/CF-TR-1994-0139.
D.E. Whitney, 1969, *"Resolved Motion Rate Control of Manipulators and Human Prostheses"*, IEEE Transactions on Man-Machine Systems.
R.L. Williams II and M.A. Murphy, 1997a, "*Implementation of Telerobotic Control Architecture Including Force-Reflecting Teleoperation*", Armstrong Lab Technical Report.
R.L. Williams II and M.A. Murphy, 1997b, "*Naturally-Transitioning Rate-to-Force Control*", submitted to the ASME Journal of Dynamic Systems, Measurement, and Control.
R.L. Williams II, 1997a, "*Cartesian Control of Force-Reflecting Hand Controllers*", Conference on Applied Mechanisms and Robotics, Cincinnati, OH.
R.L. Williams II, 1997b, "*Telerobotic Control Architecture Including Force-Reflecting Teleoperation*", Final Report, AFOSR Summer Faculty Research Program, Armstrong Laboratory.
R.L. Williams II, F.W. Harrison, and D.I. Soloway, 1997, *"Shared Control of Multiple-Manipulator, Sensor-Based Telerobotic Systems"*, 1997 IEEE International Conference on Automation and Robotics, Albuquerque, NM.
R.L. Williams II, F.W. Harrison, and D.I. Soloway, 1996, *"Naturally-Transitioning Rate-to-Force Controller for Manipulators"*, 1996 IEEE International Conference on Automation and Robotics, Minneapolis, MN.
K. Willshire, F.W. Harrison, E.F. Hogge, R.L. Williams II, and D.I. Soloway, 1992, *"Results of Telerobotic Hand Controller Study Using Force Information and Rate Control"*, AIAA Paper 92-1451.

WELL-POSEDNESS FOR A CLASS OF NONLINEAR DISTRIBUTED PARAMETER MODELS WITH TIME DELAY ARISING IN ADVANCED TOXICOKINETIC MODELING

Cynthia J. Musaute
Graduate Student
Department of Mathematics

North Carolina State University
Box 8205
Raleigh, NC 27695-8205

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

WELL-POSEDNESS FOR A CLASS OF NONLINEAR DISTRIBUTED PARAMETER MODELS WITH TIME DELAY ARISING IN ADVANCED TOXICOKINETIC MODELING

Cynthia J. Musante
Department of Mathematics
North Carolina State University

Abstract

Well-posedness of solutions was studied for a mathematical model to describe the uptake and elimination of TCDD in humans. The model includes spatial dispersion in the critical organ (the liver), time delays in tissue response, and nonlinear chemical kinetics within cells. We present a global existence and uniqueness result for a class of abstract nonlinear parabolic delay systems and discuss well-posedness for the TCDD model.

WELL-POSEDNESS FOR A CLASS OF NONLINEAR DISTRIBUTED PARAMETER MODELS WITH TIME DELAY ARISING IN ADVANCED TOXICOKINETIC MODELING

Cynthia J. Musante

1 Introduction. In this report we present existence and uniqueness results for a class of abstract nonlinear parabolic equations with time delay arising in advanced toxicokinetic modeling and address the issues involved in establishing well-posedness for a mathematical model of the human uptake and elimination of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). These efforts are basic to our eventual goal of developing computational methodologies for numerical simulation, model verification, and parameter identification for the TCDD model.

TCDD is considered to be the most toxic of a class of lipophilic chemicals which exert their effects via interaction with an intracellular protein known as the aryl hydrocarbon (Ah) receptor [B]. These chemicals are of global social concern due to their presence and persistence in the environment and their ability to produce a wide range of toxic effects in laboratory animals, including certain types of cancer [SDG, V]. Moreover, TCDD was present in Agent Orange, the herbicide used by US forces in Vietnam.

The liver appears to be a major target organ for chronic toxicity of TCDD in rodents [SDG, V]. Physiologically-based pharmaco- and toxicokinetic models which have attempted to describe the uptake, distribution, and elimination of chemicals in the body have generally used the "well-stirred" or "venous-equilibrium" model to describe events occurring in the liver. The basic assumption of this model, that the concentration of solute is uniform throughout the length of the liver acinus, does not describe the elimination of solutes with decreasing concentration gradients along the acinus following a bolus input. Moreover, the "well-stirred" model cannot accommodate spatial variations in biologically- and/or physiologically-based parameters, such as variations in enzyme activity and hepatic cell permeability. In 1985 a physiologically-based and spatially-dependent dispersion model for hepatic elimination was introduced by Roberts and Rowland [RR1]. We show in [BMT] that this model can be adapted to describe the hepatic uptake and elimination of TCDD, incorporating the complex microcirculation of the human liver as well as TCDD interaction with two intracellular proteins: the non-inducible high-affinity, low capacity Ah receptor, and an inducible low-affinity, high-capacity microsomal protein, cytochrome P450 IA2 (CYP1A2) [PG, PTG, VA]. The model is "advanced" in that it includes spatial dispersion in the critical organ (the liver), time delays in tissue response, and nonlinear chemical kinetics within cells.

The organization of this report is as follows. We begin in Section 2 with a brief discussion of the mathematical model. In Section 3 we present a general result for nonlinear parabolic systems with delay, and in Section 4 we address the issues involved with establishing well-posedness for the TCDD model. Finally, Section 5 contains a summary of our work.

2 The mathematical model. In [BMT] we present a mathematical model to describe the hepatic uptake and elimination of TCDD. Transport of solute in the liver sinusoidal (blood) region is described by a convection-dispersion equation [RR1] to account for transport via bulk flow and turbulent diffusion. The general mass transport equation is coupled with TCDD-specific equations, including the kinetics of TCDD-binding to the Ah receptor protein, induction of CYP1A2 and TCDD-binding to CYP1A2. The induction of CYP1A2 is based upon the fractional occupancy of the Ah receptor at a previous time to allow for the many intracellular processes which must occur before this Ah receptor-mediated activity is realized. Uptake of TCDD by the hepatic tissue is modeled as passive diffusion across the cell membrane, with unbound TCDD the diffusing species. Metabolism, which is considered a detoxifying step for TCDD [SDG], is modeled as a first-order process.

In order to utilize available time data on human TCDD venous blood concentrations, a simple "well-mixed" venous blood compartment was added. The model includes a sink term to describe the rate of TCDD uptake by adipose tissue, a primary storage site for TCDD in the body [SDG, V]. The resulting mathematical model is a nonlinear system of seven coupled partial and ordinary differential equations with time delay.

Boundary conditions were identified which are consistent with the assumptions of the mathematical model and yield a well-posed initial-boundary-value problem (see Appendix 1 and 2). The exit boundary condition contains an unknown boundary term and thus must be estimated; that is, we must be given observations of the solution to be used in an inverse algorithm to estimate this and possibly other unknown parameters.

Motivated by this example, in Section 3 we present a general result for abstract nonlinear parabolic systems with time delay.

3 Well-posedness for abstract nonlinear parabolic systems with time delay. The arguments for existence and uniqueness detailed in Sections 3.1- 3.5 closely follow the work of Banks, Gilliam, and Shubov [BGS] for nonlinear hyperbolic systems and are repeated here with modification for nonlinear parabolic systems. After establishing existence and uniqueness of solutions for a class of nonlinear parabolic systems on a finite time interval, and then show how this solution can be extended to systems with time delay.

3.1 Formulation of the problem. This section is concerned with establishing global existence of weak solutions for a class of abstract nonlinear parabolic systems evolving in a real separable Hilbert space $\mathcal{H}$ and holding in $\mathcal{V}^*$ as explained below:

$$\dot{y}(t) + \mathcal{A}y(t) + g(y(t)) = F(t) \text{ in } \mathcal{V}^* \tag{1}$$

$$y(0) = y_0 \tag{2}$$

for $T \in [0, T]$, $T < \infty$. Throughout we assume there is a sequence of separable Hilbert spaces $\mathcal{V}, \mathcal{H}, \mathcal{V}^*$ forming a Gelfand triple [Wl] satisfying

$$\mathcal{V} \hookrightarrow \mathcal{H} \equiv \mathcal{H}^* \hookrightarrow \mathcal{V}^*,$$

where we assume that the embedding $\mathcal{V} \hookrightarrow \mathcal{H}$ is dense and continuous with

$$|\phi|_h \leq k|\phi|_{\mathcal{V}} \text{ for } \phi \in \mathcal{V}. \tag{3}$$

We denote by $\langle \cdot, \cdot \rangle_{\mathcal{V}^*, \mathcal{V}}$ the usual duality product [Wl, BIW, BSW], which is the extension by continuity of the inner product in $\mathcal{H}$, denoted $\langle \cdot, \cdot \rangle$ throughout the remainder of this section. The norm in $\mathcal{H}$ will be denoted $|\cdot|$. The operator $\mathcal{A}$ is defined (under the assumptions below) in terms of its sesquilinear form $\sigma : \mathcal{V} \times \mathcal{V} \to R$; that is, $\mathcal{A} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ and $\langle \mathcal{A}\phi, \psi \rangle_{\mathcal{V}^*, \mathcal{V}} = \sigma(\phi, \psi)$. We make the following standing assumptions:

A1) The form σ is $\mathcal{V}$-bounded: for $\phi, \psi \in \mathcal{V}$

$$|\sigma(\phi, \psi)| \leq c_1 |\phi|_{\mathcal{V}} |\psi|_{\mathcal{V}}. \tag{4}$$

A2) The form σ is strictly coercive on $\mathcal{V}$: for $\phi \in \mathcal{V}$

$$\sigma(\phi, \phi) \geq k_1 |\phi|^2_{\mathcal{V}}, \quad k_1 > 0. \tag{5}$$

A3) The forcing term F satisfies

$$F \in L_2((0, T), \mathcal{V}^*). \tag{6}$$

A4) The nonlinear function g is a continuous nonlinear mapping from $\mathcal{H}$ into $\mathcal{H}$ satisfying

$$|g(\phi)| \leq \gamma |\phi|, \quad \phi \in \mathcal{H}, \tag{7}$$

for some constant γ.

A5) For any $\phi, \psi \in \mathcal{V}$,

$$\langle g(\phi) - g(\psi), \phi - \psi \rangle + k_1 k^{-1} |\phi - \psi|^2 \geq 0 \tag{8}$$

where k and k_1 are the constants in (3) and (5).

Condition A6) is necessary for uniqueness of solutions.

A6) For any $\phi \in \mathcal{H}$ the Fréchet derivative of g exists and satisfies

$$g'(\phi) \in \mathcal{L}(\mathcal{H}, \mathcal{H}) \text{ with } |g'(\phi)|_{\mathcal{L}(\mathcal{H},\mathcal{H})} \leq \tilde{C}_3. \tag{9}$$

Given the above hypotheses and considerations, we consider the weak or variational form of our system given by

$$\begin{gathered} \langle \dot{y}(t), \phi \rangle + \sigma(y(t), \phi) + \langle g(y(t)), \phi \rangle = \langle F(t), \phi \rangle \\ y(0) = y_0, \end{gathered} \tag{10}$$

for any $\phi \in \mathcal{V}$. We note that (1)-(2) and (10) are equivalent equations if we interpret the inner product $\langle \cdot, \cdot \rangle$ as $\langle \cdot, \cdot \rangle_{\mathcal{V}^*, \mathcal{V}}$.

3.2 The main a priori estimate. Choosing $\phi = y(t)$ in Equation (10), if a solution exists it must satisfy:

$$\frac{1}{2} \frac{d}{dt} \{|y(t)|^2\} + \sigma(y(t), y(t)) + \langle g(y(t)), y(t) \rangle = \langle F(t), y(t) \rangle_{\mathcal{V}^*, \mathcal{V}}. \tag{11}$$

Integrating from 0 to t, where $t \in [0, T]$, and using condition (5) we obtain

$$\frac{1}{2}|y(t)|^2 + k_1 \int_0^t \{|y(\xi)|_{\mathcal{V}}^2 + \langle g(y(\xi)), y(\xi) \rangle\} d\xi \leq \frac{1}{2}|y(0)|^2 + \int_0^t \langle F(\xi), y(\xi) \rangle_{\mathcal{V}^*, \mathcal{V}} d\xi.$$

Using the Cauchy-Schwarz inequality, the relation $2ab \leq a^2 + b^2$, and (7) yields

$$|y(t)|^2 + k_1 \int_0^t |y(\xi)|_{\mathcal{V}}^2 d\xi \leq |y_0|^2 + \frac{1}{k_1} \int_0^t |F(\xi)|_{\mathcal{V}^*}^2 d\xi + 2\gamma \int_0^t |y(\xi)|^2 d\xi. \tag{12}$$

By ignoring the second term on the left of (12) and applying Gronwall's lemma we obtain

$$|y(t)|^2 \leq (|y_0|^2 + \frac{1}{k_1} \int_0^T |F(\xi)|_{\mathcal{V}^*}^2 d\xi) e^{2\gamma t}. \tag{13}$$

Next, substituting (13) back into (12) we have the desired a priori bound

$$|y(t)|^2 + k_1 \int_0^t |y(\xi)|_{\mathcal{V}}^2 \leq C, \tag{14}$$

where $C = C(|y_0|, |F|_{L_2((0,T),\mathcal{V}^*)}, \gamma, k_1) < \infty$ is constant.

3.3 Galerkin approximations. Let $\{\psi_k\}_{k=1}^{\infty} \subset \mathcal{V}$ be a linearly independent total subset of $\mathcal{V}$. We define the "Galerkin" approximations for (1) by

$$y^N(t) = \sum_{k=1}^{N} c_k^N(t)\psi_k \tag{15}$$

where the coefficients $\{c_k^N(t)\}_{k=1}^N$ are chosen so that $y^N(t)$ is the unique solution of the N-dimensional system

$$\langle \dot{y}^N(t), \psi_j\rangle + \sigma(y^N(t), \psi_j) + \langle g(y^N(t)), \psi_j\rangle = \langle F(t), \psi_j\rangle_{\mathcal{V}^*,\mathcal{V}} \tag{16}$$

for $j = 1, \ldots, N$, satisfying the initial condition

$$c_k^N(0) = c_{0k}^N.$$

The set $\{c_{0k}^N\}$ is chosen such that

$$y_0 = \lim_{N\to\infty} \sum_{1}^{N} c_{0k}^N \psi_k \text{ in } \mathcal{V}.$$

Multiplying (16) by $c_j^N(t)$ and summing over $j = 1, \ldots, N$ we obtain (11) with y replaced by y^N :

$$\frac{1}{2}\frac{d}{dt}\{|y^N(t)|^2\} + \sigma(y^N(t), y^N(t)) + \langle g(y^N(t)), y^N(t)\rangle = \langle F(t), y^N(t)\rangle_{\mathcal{V}^*,\mathcal{V}}.$$

Repeating the above arguments, we then obtain

$$|y^N(t)|^2 + k_1 \int_0^t |y^N(\xi)|_{\mathcal{V}}^2 d\xi \le \tilde{C}, \tag{17}$$

where the constant $\tilde{C}$ is independent of N, depending on y_0 and F as in the constant C of (14). The convergence of $y_0^N \to y_0$ in $\mathcal{V}$ guarantees uniform boundedness of $|y_0^N|_{\mathcal{V}}$.

3.4 Convergence of the Galerkin approximations. To establish existence of solutions to (1)-(2) we will use the bounds of (17) to extract successive subsequences of the Galerkin approximations. We show that the final subsequence converges to a solution for the problem. The subsequences selected at each step will be denoted by the same symbol $\{y^N\}$.

It follows from (17) that the set $\{y^N\}$ is bounded in $C([0,T],\mathcal{V}) \subset L_2((0,T),\mathcal{V})$. Hence we may conclude that there exists a subsequence such that

$$y^N \xrightarrow{w} y \text{ (weakly) in } L_2((0,T),\mathcal{V}). \tag{18}$$

Moreover, (17) also implies

$$y^N(t) \xrightarrow{w} y(t) \text{ (weakly) in } \mathcal{H} \text{ for each } t \in [0,T]. \tag{19}$$

The following lemma is needed to carry out the proof of existence.

Lemma 1 *There exists a subsequence* $\{y^N\}$ *of the original sequence of Galerkin approximations and* $y \in L_2((0,T),\mathcal{V})$ *such that*

a)

$$y^N \xrightarrow{w} y \text{ (weakly) in } L_2((0,T),\mathcal{V}) \tag{20}$$

b)

$$y^N(t) \xrightarrow{w} y(t) \text{ (weakly) in } \mathcal{H} \text{ for each } t \in [0,T] \tag{21}$$

c) *There exists* $h \in L_2((0,T),\mathcal{H})$ *such that*

$$g(y^N) \xrightarrow{w} h \text{ (weakly) in } L_2((0,T),\mathcal{H}). \tag{22}$$

Proof. The statements (20) and (21) are repetitions of (18) and (19). Statement c) follows immediately from (3), (7), and (17).

3.5 Existence of weak solutions. In this section we obtain the fundamental existence result by letting $N \to \infty$ in an integral identity for the Galerkin approximations.

Theorem 1 *Under assumptions* A1)-A5) *there exists a weak solution* y *of* (1)-(2) *with* $y \in L_2((0,T),\mathcal{V})$ *and* $\dot{y} \in L_2((0,T),\mathcal{V}^*)$. *Furthermore, if condition* A6) *holds, the solution is unique.*

Proof. We define $\mathcal{L}_T = L_2((0,T),\mathcal{V})$ and denote by P_M $(M = 1,2,\ldots)$ the class of functions $\eta \in \mathcal{L}_T$ which can be represented in the form

$$\eta(t) = \sum_{k=1}^{M} a_k(t)\psi_k, \tag{23}$$

where $a_k \in C^1([0,T])$. Let

$$P = \bigcup_{M=1}^{\infty} P_M. \tag{24}$$

P is dense in $\mathcal{L}_T$. We multiply the j^{th} equation in (16) by $a_j(t)$, take the sum from 1 to M and integrate over $[0,t]$ (by parts in the first term) to obtain

$$\begin{aligned}\int_0^t \{-\langle y^N(\xi),\dot{\eta}(\xi)\rangle + \sigma(y^N(\xi),\eta(\xi)) + \langle g(y^N(\xi)),\eta(\xi)\rangle\}d\xi + \langle y^N(t),\eta(t)\rangle \\ = \langle y^N(0),\eta(0)\rangle + \textstyle\int_0^t \langle F(\xi),\eta(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi,\end{aligned} \tag{25}$$

which is satisfied for all $\eta \in P_M$, $M \le N$.

We fix $\eta \in P_M$ with $M \leq N$, and use (20), (21), and (22). Passing to the limit $N \to \infty$ in (25) we obtain

$$\int_0^t \{-\langle y(\xi), \dot{\eta}(\xi)\rangle + \sigma(y(\xi), \eta(\xi)) + \langle h(\xi), \eta(\xi)\rangle\} d\xi + \langle y(t), \eta(t)\rangle$$
$$= \langle y_0, \eta(0)\rangle + \int_0^t \langle F(\xi), \eta(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi. \tag{26}$$

Note all the statements of Lemma 1 are true for any interval $[0, t]$, for $t \leq T$. To pass to the limit in the first term under the integral sign we only need the weak convergence $y^N \xrightarrow{w} y$ in $L_2((0,T), \mathcal{H})$ as in (20). In the second term we note that for fixed $\eta \in \mathcal{V}$ the mapping $y \to \int_0^t \sigma(y(\xi), \eta(\xi)) d\xi$ is a bounded linear functional on $L_2([0,t], \mathcal{V})$ due to (4), and therefore this functional is weakly continuous. This allows us to pass to the limit due to (20). In the third term we can pass to the limit due to (22). In the term outside the integral in the left side of (25) we can pass to the limit due to (21). We use the fact that $y^N(0) \to y_0$ strongly in $\mathcal{V}$ and therefore in $\mathcal{H}$ as $N \to \infty$ in the first term on the right hand side of (25).

Now for each j choose $\eta_j(t) = a(t)\psi_j$ where we further restrict a so that $a \in C_0^\infty[0,T]$. We obtain from (26)

$$\int_0^T \dot{a}(\xi)\langle -y(\xi), \psi_j\rangle d\xi + \int_0^T a(\xi)\{\sigma(y(\xi), \psi_j) + \langle h(\xi), \psi_j\rangle - \langle F(\xi), \psi_j\rangle_{\mathcal{V}^*,\mathcal{V}}\} d\xi = 0.$$

This implies

$$\langle \dot{y}(t), \psi_j\rangle + \sigma(y(t), \psi_j) + \langle h(t), \psi_j\rangle = \langle F(t), \psi_j\rangle_{\mathcal{V}^*,\mathcal{V}} \tag{27}$$

for each j and $t \in [0,T]$. Since $\{\psi_j\}$ is total in $\mathcal{V}$ we have $\dot{y} \in L_2((0,T), \mathcal{V}^*)$ and for all $\phi \in \mathcal{V}$

$$\langle \dot{y}(t), \phi\rangle + \sigma(y(t), \phi) + \langle h(t), \phi\rangle = \langle F(t), \phi\rangle_{\mathcal{V}^*,\mathcal{V}} \tag{28}$$

which except for the term involving the limit function h is the equation for weak solutions (10). We show first that $y(0) = y_0$ and then argue that the h term under the integral of (26) can be replaced by $g(y)$ which yields that the limit function y is a weak solution.

To show $y(0) = y_0$ we return to Equation (26) which holds for each $\eta_j = a(t)\psi_j$, where $a \in C^1[0,T]$ and $a(0) \neq 0$. Integrating by parts in the first term in (26) and using (27) we obtain

$$\langle y(t), \eta(t)\rangle + \langle -y(\xi), \eta(\xi)\rangle_{\xi=0}^{\xi=t} = \langle y_0, \eta(0)\rangle,$$

or

$$\langle y(0), \psi_j\rangle a(0) = \langle y_0, \psi_j\rangle a(0),$$

for all j, from which it follows that $y(0) = y_0$.

To prove that the limit function h is the correct term we use the Minty-Browder monotonicity method, as in [BGS].

Lemma 2 *For any $\eta \in \mathcal{L}_T$ and for $t \in [0,T]$*

$$\int_0^t \langle g(y(\xi)), \eta(\xi)\rangle d\xi = \int_0^t \langle h(\xi), \eta(\xi)\rangle d\xi.$$

Proof. From (8) and (3) we obtain

$$\int_0^t \{\langle g(u(\xi)) - g(v(\xi)), u(\xi) - v(\xi)\rangle + k_1 |u(\xi) - v(\xi)|_{\mathcal{V}}^2\} d\xi \geq 0 \tag{29}$$

for any $u, v \in \mathcal{L}_T$.

Select $u = y^N \in \mathcal{L}_T$ and any $v \in P^M \subset \mathcal{L}_T$ with $M \leq N$. Using condition (5), we obtain

$$\int_0^t \langle g(y^N(\xi)) - g(v(\xi)), y^N(\xi) - v(\xi)\rangle d\xi + \int_0^t \sigma(y^N(\xi) - v(\xi), y^N(\xi) - v(\xi)) d\xi \geq 0. \tag{30}$$

Returning to (25), we take $\eta = y^N - v$ which is permissible since $y^N \in P^N$ and $v \in P^M$, $M \leq N$. This yields

$$\int_0^t \langle g(y^N(\xi)), y^N(\xi) - v(\xi)\rangle d\xi = \int_0^t \langle y^N(\xi), \dot{y}^N(\xi) - \dot{v}(\xi)\rangle d\xi - \int_0^t \sigma(y^N(\xi), y^N(\xi) - v(\xi)) d\xi$$
$$- \langle y^N(t), y^N(t) - v(t)\rangle + \langle y^N(0), y^N(0) - v(0)\rangle + \int_0^t \langle F(\xi), y^N(\xi) - v(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi. \tag{31}$$

Now substituting (31) into (30) we obtain for all $v \in P_M$, $M \leq N$

$$-\frac{1}{2}|y^N(t)|^2 + \frac{1}{2}|y^N(0)|^2 + \int_0^t \{-\langle y^N(\xi), \dot{v}(\xi)\rangle - \sigma(v(\xi), y^N(\xi) - v(\xi)) - \langle g(v(\xi)), y^N(\xi) - v(\xi)\rangle\} d\xi$$
$$+ \int_0^t \langle F(\xi), y^N(\xi) - v(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi + \langle y^N(t), v(t)\rangle - \langle y^N(0), v(0)\rangle \geq 0. \tag{32}$$

We can pass to the limit $N \to \infty$ in (32) to obtain

$$-\frac{1}{2}|y(t)|^2 + \frac{1}{2}|y_0|^2 + \int_0^t \{-\langle y(\xi), \dot{v}(\xi)\rangle - \sigma(v(\xi), y(\xi) - v(\xi)) - \langle g(v(\xi)), y(\xi) - v(\xi)\rangle\} d\xi$$
$$+ \int_0^t \langle F(\xi), y(\xi) - v(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi + \langle y(t), v(t)\rangle - \langle y_0, v(0)\rangle \geq 0. \tag{33}$$

We can pass to the limit in all terms under the integral sign due to the weak convergence results presented in Lemma 1. In the first term after the integral sign we can pass to the limit due to the weak convergence of $y^N(t) \xrightarrow{w} y(t)$ in $\mathcal{H}$ for each $t \in [0,T]$. Strong convergence of $y^N(0) \to y_0$ in $\mathcal{V}$ implies strong convergence

in $\mathcal{H}$, allowing us to pass to the limit in the second and last terms in the inequality. The first term in (32) requires further comment, as we only have weak convergence $y^N(t) \xrightarrow{w} y(t)$ in $\mathcal{H}$ for any $t \in [0,T]$. However, since norms are weakly lower semicontinuous in Hilbert spaces, we have

$$|y(t)|^2 \leq \liminf_{N\to\infty} |y^N(t)|^2.$$

Therefore, we can pass to the limit in (32), obtaining the desired result (33). Note that (33) is valid for any $v \in P = \cup_{M=1}^{\infty} P^M$ and therefore for all $v \in \mathcal{L}_T$. Now we return to (26) with $\eta = -v$ where v is from (33):

$$\int_0^t \{\langle y(\xi), \dot{v}(\xi)\rangle - \sigma(y(\xi), v(\xi)) - \langle h(\xi), v(\xi)\rangle\} d\xi - \langle y(t), v(t)\rangle + \langle y_0, v(0)\rangle + \int_0^t \langle F(\xi), v(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi = 0. \quad (34)$$

Also, set $\eta = y$ in (26) to obtain

$$\frac{1}{2}(|y(t)|^2 - |y_0|^2) + \int_0^t \{\sigma(y(\xi), y(\xi)) + \langle h(\xi), y(\xi)\rangle\} d\xi - \int_0^t \langle F(\xi), y(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} d\xi = 0. \quad (35)$$

Adding inequality (33) and the relations (34) and (35) yields

$$\int_0^t \{\sigma(y(\xi) - v(\xi), y(\xi) - v(\xi)) + \langle h(\xi) - g(v(\xi)), y(\xi) - v(\xi)\rangle\} d\xi \geq 0. \quad (36)$$

For any $\theta > 0$ and $\omega \in \mathcal{L}_T$, choose

$$v(t) = y(t) - \theta\omega(t). \quad (37)$$

Substituting (37) into (36) and dividing by $\theta > 0$ we obtain

$$\int_0^t \{\theta\sigma(\omega(\xi), \omega(\xi)) + \langle h(\xi) - g(y(\xi) - \theta\omega(\xi)), \omega(\xi)\rangle\} d\xi \geq 0 \quad (38)$$

for any $\omega \in \mathcal{V}, \theta > 0$. In (38) we can pass to the limit as $\theta \to 0$ which yields the inequality

$$\int_0^t \langle h(\xi) - g(y(\xi)), \omega(\xi)\rangle d\xi \geq 0 \quad (39)$$

using the fact that $g : \mathcal{H} \to \mathcal{H}$ is a continuous mapping.

The inequality (39) holds for all $\omega \in \mathcal{L}_T$ only if it holds for equality. Suppose that for some $\omega \in \mathcal{L}_T$ we have a strict inequality in (39). Then replacing ω by $-\omega$ we obtain a contradiction to (39). Thus we have

$$\int_0^t \langle h(\xi) - g(y(\xi)), \omega(\xi)\rangle d\xi = 0 \text{ for all } \omega \in \mathcal{L}_T. \quad (40)$$

Therefore, Lemma 2 is established and the proof of existence is complete.

We now address the question of uniqueness of solutions. Let w, v be two solutions of (10) corresponding to the initial condition y_0 and forcing term F. Note that $u \equiv w - v$ satisfies $u(0) = 0$ and

$$\langle \dot{u}(t), \eta\rangle_{\mathcal{V}^*,\mathcal{V}} + \sigma(u(t), \eta) + \langle g(w(t)) - g(v(t)), \eta\rangle = 0 \text{ for all } \eta \in \mathcal{V}, t \in [0,T]. \quad (41)$$

Observe (41) as well as (10) holds for all $\eta \in \mathcal{L}_T$. For fixed $t \in [0,T]$ choose $\eta = u$ in (41) and integrate form 0 to t :

$$\int_0^t \{\langle \dot{u}(\xi), u(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}} + \sigma(u(\xi), u(\xi)) + \langle \Delta g(\xi), u(\xi)\rangle\} d\xi = 0$$

where $\Delta g(\xi) \equiv g(w(\xi)) - g(v(\xi))$. Since $\langle \dot{u}(\xi), u(\xi)\rangle = \frac{1}{2}\frac{d}{d\xi}\{|u(\xi)|^2\}$, this implies

$$|u(t)|^2 + 2\int_0^t \{\sigma(u(\xi), u(\xi)) + \langle \Delta g(\xi), u(\xi)\rangle\} d\xi = 0.$$

From condition (5) and (9) we obtain

$$\begin{aligned} |u(t)|^2 + 2k_1 \int_0^t |u(\xi)|_{\mathcal{V}}^2 d\xi &\leq 2\Big| \int_0^t \langle \Delta g(\xi), u(\xi)\rangle d\xi \Big| \\ &\leq 2\Big| \int_0^t \langle \int_0^1 g'(\theta w(\xi) + (1-\theta)v(\xi))[w(\xi) - v(\xi)] d\theta, u(\xi)\rangle d\xi \Big| \\ &\leq 2\int_0^t \tilde{C}_3 |u(\xi)|^2 d\xi. \end{aligned}$$

Ignoring the second term in the inequality, we have for any $t \in [0,T]$

$$|u(t)|^2 \leq 2\int_0^t \tilde{C}_3 |u(\xi)|^2 d\xi.$$

This implies $u \equiv 0$ on any finite interval by Gronwall's lemma. Therefore, Theorem 1 is proven.

3.6 Weakening condition A5) on the nonlinearity. In actual fact, for $F \in L_2((0,T),\mathcal{H})$, one can eliminate the "convexity condition" on the nonlinearity (condition A5)) by strengthening the convergence of the Galerkin approximations.

Multiplying (16) by $\dot{c}_j^N(t)$ and summing over $j = 1, \ldots, N$ we obtain

$$|\dot{y}^N(t)|^2 + \frac{d}{dt}\sigma(y^N(t), y^N(t)) + \langle g(y^N(t)), \dot{y}^N(t)\rangle = \langle F(t), \dot{y}^N(t)\rangle. \tag{42}$$

We now repeat the arguments in Section 3.2. Integrating from 0 to t for $t \in [0,T]$, yields

$$\int_0^t |\dot{y}^N(\xi)|^2 d\xi + \sigma(y^N(t), y^N(t)) + \int_0^t \langle g(y^N(\xi)), \dot{y}^N(\xi)\rangle d\xi = \sigma(y^N(0), y^N(0)) + \int_0^t \langle F(\xi), \dot{y}^N(\xi)\rangle d\xi.$$

By conditions (4) and (5) we have

$$k_1 |y^N(t)|_{\mathcal{V}}^2 + \int_0^t |\dot{y}^N(\xi)|^2 d\xi \leq c_1 |y^N(0)|_{\mathcal{V}}^2 + \int_0^t \langle F(\xi), \dot{y}^N(\xi)\rangle d\xi + \int_0^t |\langle g(y^N(\xi)), \dot{y}^N(\xi)\rangle| d\xi.$$

Using the Cauchy-Schwarz inequality, the assumption that F is in $L_2((0,T),\mathcal{H})$ and condition (7) we have for all $\epsilon > 0$

$$k_1 |y^N(t)|_{\mathcal{V}}^2 + (1-\epsilon)\int_0^t |\dot{y}^N(\xi)|^2 d\xi \leq c_1 |y^N(0)|_{\mathcal{V}}^2 + \frac{1}{2\epsilon}|F|_{L_2((0,T),\mathcal{H})}^2 + \frac{\gamma^2}{2\epsilon}|y^N|_{L_2((0,T),\mathcal{H})}^2.$$

The convergence of $y^N(0) \to y_0$ strongly in $\mathcal{V}$ guarantees uniform boundedness of $|y^N(0)|_{\mathcal{V}}$. Furthermore, $\{y^N\}$ is uniformly bounded in $L_2((0,T),\mathcal{V}) \subset L_2((0,T),\mathcal{H})$ by a constant $\tilde{C}$ (17). Therefore, choosing $\epsilon > 0$ such that $1-\epsilon > 0$ we obtain

$$k_1|y^N(t)|^2_{\mathcal{V}} + (1-\epsilon)\int_0^t |\dot{y}^N(\xi)|^2 d\xi \leq \hat{C}, \tag{43}$$

where the constant $\hat{C}$ is independent of N, depending only on y_0, F, γ, k_1 and $\tilde{C}$. Hence, $\{y^N\}$ and $\{\dot{y}^N\}$ are uniformly bounded in $C([0,T],\mathcal{V})$ and $L_2((0,T),\mathcal{H})$, respectively.

Lemma 3 *For $F \in L_2((0,T),\mathcal{H})$, there exists a subsequence of the original sequence of Galerkin approximations such that*

$$\{y^N\} \to y \text{ strongly in } C([0,T],\mathcal{H}).$$

Proof. We use the following version of the Arzela-Ascoli theorem [NS, Theorem 3.17.24]: If Y is a complete metric space and $\mathcal{F} \subset C([0,T],Y)$, then $\mathcal{F}$ is relatively compact if and only if $\mathcal{F}$ is equicontinuous and $\{f(t) : f \in \mathcal{F}\}$ is relatively compact in Y for each $t \in [0,T]$.

Letting $Y = \mathcal{H}$ and $\mathcal{F} = \{y^N\} \subset C([0,T],\mathcal{H})$ in the statement of the theorem, it can be shown that the subset $\{y^N\}$ is equicontinuous in $C([0,T],\mathcal{H})$. Recall for each $t \in [0,T]$, $\{y^N(t)\}$ is uniformly bounded in $\mathcal{V}$. Furthermore, since $\mathcal{V}$ embeds compactly in $\mathcal{H}$, for each $t \in [0,T]$ the set $\{y^N(t)\}$ is relatively compact in $\mathcal{H}$. Thus, the Arzela-Ascoli theorem allows us to conclude that $\{y^N\}$ is relatively compact in $C([0,T],\mathcal{H})$. Hence, there exists a subsequence of the original Galerkin approximations such that $y^N \to y$ strongly in $C([0,T],\mathcal{H})$.

Theorem 2 *Under the assumptions* A1)-A3) *with $F \in L_2((0,T),\mathcal{H})$, there exists a weak solution y of* (1)-(2) *with $y \in L_2((0,T),\mathcal{V})$ and $\dot{y} \in L_2((0,T),\mathcal{V}^*)$. Furthermore, if condition* (9) *holds, the solution is unique.*

Proof. We need only show that we can pass to the limit in (25). Lemma 3 and the fact that g is a continuous mapping from $\mathcal{H}$ into $\mathcal{H}$ (7) guarantee this result. Thus, (26) becomes

$$\int_0^t \{-\langle y(\xi),\dot{\eta}(\xi)\rangle + \sigma(y(\xi),\eta(\xi)) + \langle g(y(\xi)),\eta(\xi)\rangle\}\, d\xi + \langle y(t),\eta(t)\rangle = \langle y_0,\eta(0)\rangle + \int_0^t \langle F(\xi),\eta(\xi)\rangle_{\mathcal{V}^*,\mathcal{V}}\, d\xi. \tag{44}$$

Similar arguments to the proof of Theorem 1 follow from (44).

3.7 Existence of weak solutions for systems with time delay. This section is concerned with establishing existence of weak solutions for a class of nonlinear parabolic problems with time delay τ, $0 < \tau < \infty$:

$$\dot{y}(t) + \mathcal{A}y(t) + \mathcal{A}_D(y(t-\tau)) + g(y(t)) + g_D(y(t-\tau)) = F(t) \text{ in } \mathcal{V}^*, \tag{45}$$

$$y(0) = y_0, \tag{46}$$

where the linear operator $\mathcal{A}_D$ is defined in terms of its sesquilinear form $\sigma_D : \mathcal{V} \times \mathcal{V} \to R$. That is, $\mathcal{A}_D \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ and $\langle \mathcal{A}_D\phi, \psi\rangle_{\mathcal{V}^*,\mathcal{V}} = \sigma_D(\phi, \psi)$. Furthermore, we consider the same setting as in Section 3.1; we assume A1)-A6) and

D6) The form σ_D is $\mathcal{V}$-bounded: for $\phi, \psi \in \mathcal{V}$,

$$|\sigma_D(\phi, \psi)| \leq c_2|\phi|_{\mathcal{V}}|\psi|_{\mathcal{V}}. \tag{47}$$

D7) The function g_D is a continuous nonlinear mapping from $\mathcal{H}$ into $\mathcal{H}$.

Theorem 3 *Under assumptions* A1)-A5) *and* D6)-D7) *there exists a weak solution y of* (45)-(46) *with $y \in L_2([0,T], \mathcal{V})$ and $\dot{y} \in L_2([0,T], \mathcal{V}^*)$. Furthermore, if condition* A6) *holds, the solution is unique.*

Remark 1 The terms involving $\mathcal{A}_D$ and g_D in (45) do not affect the solution on the interval $[0, \tau)$. Therefore, by Theorem 1 we have the existence of a unique solution y of the equation

$$\langle \dot{y}(t), \phi\rangle_{\mathcal{V}^*,\mathcal{V}} + \sigma(y(t), \phi) + \langle g(y(t)), \phi\rangle = \langle F(t), \phi\rangle_{\mathcal{V}^*,\mathcal{V}}$$

$$y(0) = y_0$$

for $t \in (0, \tau)$ and $\phi \in \mathcal{V}$. We set

$$\Phi(t) = y(t-\tau), \qquad \text{for } t \in [\tau, 2\tau),$$

and note that Φ is a known function of time on the interval $[\tau, 2\tau)$. Defining $G \in L_2((\tau, 2\tau), \mathcal{V}^*)$ by

$$G(t)(\phi) = \langle \mathcal{A}_D\Phi(t) + g_D(\Phi(t)), \phi\rangle$$

for $\phi \in \mathcal{V}$ and $t \in [\tau, 2\tau)$, we then seek a solution of the equation

$$\langle \dot{y}(t), \phi\rangle_{\mathcal{V}^*,\mathcal{V}} + \sigma(y(t), \phi) + \langle g(y(t)), \phi\rangle = \langle F(t) - G(t), \phi\rangle_{\mathcal{V}^*,\mathcal{V}} \tag{48}$$

for all $\phi \in \mathcal{V}$ and $t \in [\tau, 2\tau)$. By Theorem 1 a solution y of (48) exists and is unique on the interval $[\tau, 2\tau)$. We then continue in this manner to "step" the solution forward by time intervals of length τ until the desired final time T. This is the standard "method of steps" or "method of continuation" used in delay differential equations by which the solution is extended forward in time [BC].

4 Well-posedness for the TCDD dispersion model. We consider the system in Appendix 1 given by Equations (50)-(58) with an appropriate set of initial conditions. To simplify our analysis, we make a change of variable from C_B to $\tilde{C}_B$ defined by

$$\tilde{C}_B(t,z) = C_B(t,z) - C_{Vb}(t).$$

Note that $\tilde{C}_B$ satisfies a homogenous Dirichlet boundary condition at $z = 0$. We then multiply each equation by a function ϕ in a "suitable" class of test functions and integrate in space over $[0, l]$, by parts in the first equation (50) only.

Unless otherwise indicated, throughout the remainder of this report $\langle \cdot, \cdot \rangle$ denotes the usual L_2 inner product:

$$\langle f, g \rangle = \int_0^l f(\xi) g(\xi) d\xi.$$

4.1 State space setting and problem formulation. We take $V = H_L^1(0,l)$ and $H = L_2(0,l)$, where

$$H_L^1(0,l) = \{\phi \in H^1(0,l) | \phi(0) = 0\}.$$

For all $\phi \in V$, define $|\phi|_V = |\phi'|_H$. Note V and H form the usual Gelfand triple as discussed in Section 3.1.

We now define the state space $\mathcal{V} = V \times H^5 \times R$ and $\mathcal{H} = H^6 \times R$. Note that under the Gelfand triple formulation above we have $\mathcal{V} \hookrightarrow \mathcal{H} \equiv \mathcal{H}^* \hookrightarrow \mathcal{V}^*$ with duality pairing $\langle \cdot, \cdot \rangle_{\mathcal{V}^*, \mathcal{V}}$, where $\mathcal{V}^* = V^* \times H^5 \times R$. The inner product in $\mathcal{H}$ is defined

$$\langle \phi, \psi \rangle_{\mathcal{H}} = \sum_{k=1}^{6} \langle \phi_k, \psi_k \rangle + \langle \phi_7, \psi_7 \rangle_R, \qquad \text{for all } \phi, \psi \in \mathcal{H}.$$

In weak form, the system (50)-(58) can be expressed

$$\langle \dot{y}(t), \phi \rangle + \sigma(y(t), \phi) + \sigma_D(y(t - \tau_c), \phi) + \langle g(y(t)), \phi \rangle + \langle g_D(y(t - \tau_r)), \phi \rangle = \langle f(t), \phi \rangle \tag{49}$$

for all $\phi \in \mathcal{V}$, where

$$\begin{aligned} f(t) &= [\dot{q}_3(t) - I(t) - a_5 q_2(t)\delta_l, 0, 0, k_{s(Ah)}, 0, k_{s(Pr)}, I(t) - \dot{q}_3(t)]^T \\ g(y) &= [0, k_{+1} y_2 y_4 + k_{+2} y_2 y_6, -k_{+1} y_2 y_4, k_{+1} y_2 y_4, -k_{+2} y_2 y_6, k_{+2} y_2 y_6, 0]^T \\ g_D(y) &= [0, 0, 0, 0, 0, (\tfrac{I_{Pr}}{V_H}) \tfrac{y_3}{y_3 + y_4}, 0]^T \end{aligned}$$

and the sesquilinear forms $\sigma, \sigma_D : \mathcal{V} \times \mathcal{V} \to R$ are defined

$$\begin{aligned}
\sigma(\psi,\phi) &= \langle a_1\psi_1' - a_2\psi_1, \phi_1'\rangle + \langle a_4\psi_1 - a_3\psi_2 + (a_4-\alpha)\psi_7, \phi_1\rangle - \langle a_2\psi_7\delta_l, \phi_1\rangle + \langle b_2\psi_2, \phi_2\rangle \\
&\quad + \langle k_{-1}\psi_3, \phi_3\rangle - \langle (b_1\psi_1 + k_{-1}\psi_3 + k_{-2}\psi_5 + b_1\psi_7), \phi_2\rangle + \langle k_{d(Ah)}\psi_4, \phi_4\rangle \\
&\quad - \langle k_{-1}\psi_3, \phi_4\rangle + \langle k_{-2}\psi_5, \phi_5\rangle + \langle k_{d(Pr)}\psi_6, \phi_6\rangle - \langle k_{-2}\psi_5, \phi_6\rangle + \langle \alpha\psi_7, \phi_7\rangle_R \\
\sigma_D(\psi,\phi) &= \alpha[\langle \psi_1(l), \phi_1\rangle + \langle \psi_7, \phi_1\rangle - \langle \psi_1(l), \phi_7\rangle_R - \langle \psi_7, \phi_7\rangle_R].
\end{aligned}$$

4.2 The TCDD dispersion model. We state the following without proof:

(H1) σ is $\mathcal{V}$-continuous: There exists a constant $\gamma_1 > 0$ such that for all $\psi, \phi \in \mathcal{V}$

$$|\sigma(\psi,\phi)| \le \gamma_1 |\psi|_{\mathcal{V}} |\phi|_{\mathcal{V}}.$$

(H2) σ is $\mathcal{V}$-coercive: There exist constants $k_1 > 0, \lambda_0 > 0$ such that for all $\phi \in \mathcal{V}$

$$\sigma(\phi,\phi) \ge k_1 |\phi|_{\mathcal{V}}^2 - \lambda_0 |\phi|_{\mathcal{H}}^2.$$

(H3) σ_D is $\mathcal{V}$-continuous: There exists a constant $\gamma_2 > 0$ such that for all $\psi, \phi \in \mathcal{V}$

$$|\sigma_D(\psi,\phi)| \le \gamma_2 |\psi|_{\mathcal{V}} |\phi|_{\mathcal{V}}.$$

Under the conditions (H1),(H3) above, there exist continuous linear operators $\mathcal{A}, \mathcal{A}_D : \mathcal{V} \to \mathcal{V}^*$ such that $\sigma(\phi,\psi) = \langle \mathcal{A}\phi, \psi\rangle_{\mathcal{V}^*,\mathcal{V}}$ and $\sigma_D(\phi,\psi) = \langle \mathcal{A}_D\phi, \psi\rangle_{\mathcal{V}^*,\mathcal{V}}$ for $\phi, \psi \in \mathcal{V}$ [Wl]. This is similar to the system in (45), with two important distinctions.

First, Theorem 3 does not apply since the "convexity condition" (condition A5)) does not hold; that is, any bounds on the nonlinear term g or it's derivative will involve the magnitudes of the reaction coefficients k_{+1} and k_{+2}. As shown in Theorem 2, it is possible to relax this condition on g provided the forcing term F is in $L_2((0,T),\mathcal{H})$. However, for the dispersion model, the forcing term involves the pointwise evaluation of a test function at the right boundary; that is, F involves a dirac delta function, δ_l, and $\delta_l \notin L_2((0,T),\mathcal{H})$.

5 Summary. Motivated by a convection-dispersion model for the uptake and elimination of TCDD in humans, we have presented a general result for existence and uniqueness of solutions for a class of abstract nonlinear parabolic equations. We have shown that one condition pertaining to the nonlinearity, the so-called "convexity condition", can be relaxed provided one has greater regularity on the forcing term. These results can be extended to a class of nonlinear parabolic systems with time delay by the "method of steps" used in

the study of delay differential equations. Our efforts continue toward the goal of establishing well-posedness for the TCDD model, which does not satisfy the conditions of the general result presented in this report.

Acknowledgments. The author is deeply grateful to Dr. Richard Albanese, AL/OES Mathematical Products Division, Brooks AFB, TX and Dr. H.T. Banks, Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC for their support, encouragement, and guidance.

Appendix 1: The model system of equations.

$$(V_B + V_D\frac{f_{u_B}}{f_{u_D}})\frac{\partial C_B}{\partial t} = V_B\mathcal{D}\frac{\partial^2 C_B}{\partial z^2} - V_B v\frac{\partial C_B}{\partial z} + P(C_{u_H} - f_{u_B}C_B) \tag{50}$$

$$\begin{aligned}\frac{\partial C_{u_H}}{\partial t} &= \frac{Pf_{u_B}}{V_H}C_B - (\frac{P}{V_H} + k_3)C_{u_H} - k_{+1}C_{u_H}C_{\text{Ah}} + k_{-1}C_{\text{Ah-TCDD}} \\ &\quad -k_{+2}C_{u_H}C_{\text{Pr}} + k_{-2}C_{\text{Pr-TCDD}}\end{aligned} \tag{51}$$

$$\frac{\partial C_{\text{Ah-TCDD}}}{\partial t} = k_{+1}C_{u_H}C_{\text{Ah}} - k_{-1}C_{\text{Ah-TCDD}} \tag{52}$$

$$\frac{\partial C_{\text{Ah}}}{\partial t} = k_{-1}C_{\text{Ah-TCDD}} - k_{+1}C_{u_H}C_{\text{Ah}} - k_{d(Ah)}C_{\text{Ah}} + k_{s(Ah)} \tag{53}$$

$$\frac{\partial C_{\text{Pr-TCDD}}}{\partial t} = k_{+2}C_{u_H}C_{\text{Pr}} - k_{-2}C_{\text{Pr-TCDD}} \tag{54}$$

$$\frac{\partial C_{\text{Pr}}}{\partial t} = k_{-2}C_{\text{Pr-TCDD}} - k_{+2}C_{u_H}C_{\text{Pr}} - k_{d(Pr)}C_{\text{Pr}} + k_{s(Pr)} \tag{55}$$

$$+(\frac{I_{Pr}}{V_H})\frac{C_{\text{Ah-TCDD}}(t-\tau_r, z)}{C_{\text{Ah}}(t-\tau_r, z) + C_{\text{Ah-TCDD}}(t-\tau_r, z)} \tag{56}$$

$$\frac{dC_{Vb}}{dt}(t) = \frac{Q_{Vb}}{V_{Vb}}(C_B(t-\tau_c, l) - C_{Vb}(t)) + I(t) - \dot{q}_3(t) \tag{57}$$

$$\begin{aligned}&C_B(t,0) = C_{Vb}(t) \\ &vC_B(t,l) - \mathcal{D}\tfrac{\partial C_B}{\partial z}(t,l) = q_2(t).\end{aligned} \tag{58}$$

Appendix 2: Notation. l denotes a unit length, m a unit mass, and t a unit time.

ABBR.	DESCRIPTION	UNITS
V_B	liver blood volume	l^3
V_D	Disse space volume	l^3
V_H	hepatocyte volume	l^3
v	average liver fluid flow velocity	l/t
C_B	total concentration of TCDD in liver blood volume	m/l^3
C_{u_H}	concentration of unbound TCDD in hepatocytes	m/l^3
C_{Ah}	total concentration of Ah receptor protein in hepatocytes	m/l^3
C_{Pr}	total concentration of CYP1A2 in hepatocytes	m/l^3
$C_{\text{Ah-TCDD}}$	concentration of Ah receptor-TCDD complex in hepatocytes	m/l^3
$C_{\text{Pr-TCDD}}$	concentration of CYP1A2-TCDD complex in hepatocytes	m/l^3
f_{u_B}	fraction of TCDD unbound in liver blood volume	
f_{u_D}	fraction of TCDD unbound in space of Disse	
P	permeability coefficient of hepatocytes to TCDD	l^3/t
$\mathcal{D}$	axial dispersion coeffecient	l^2/t
$k_{d(Ah)}$	rate constant for thermal inactivation of Ah receptor protein	$/t$
$k_{d(Pr)}$	rate constant for degradation of CYP1A2	$/t$
k_{+1}	association rate constant of TCDD and Ah receptor	$l^3/(mt)$
k_{-1}	dissociation rate constant of Ah receptor-TCDD complex	$/t$
k_{+2}	association rate constant of TCDD and CYP1A2	$l^3/(mt)$
k_{-2}	dissociation rate constant of CYP1A2-TCDD complex	$/t$
k_3	apparent first-order metabolic clearance rate of TCDD	$/t$
$k_{s(Ah)}$	rate constant for synthesis of Ah receptor protein	$m/(l^3t)$
$k_{s(Pr)}$	rate constant for basal synthesis of CYP1A2	$m/(l^3t)$
I_{Pr}	rate constant for induction of CYP1A2 in presence of TCDD	m/t
τ_r	lag time between TCDD binding to Ah receptor and response (presence of induced CYP1A2)	t
C_{Vb}	TCDD venous blood concentration	m/l^3
Q_{Vb}	venous blood volumetric flow rate	l^3/t
V_{Vb}	venous blood volume	l^3
$\dot{q}_3(t)$	rate of uptake of TCDD by fatty tissue	$m/(l^3t)$
$I(t)$	input concentration of TCDD in venous blood	m/l^3
τ_c	lag time for circulation of venous blood	t

References

[B] L.S. Birnbaum, *The mechanism of dioxin toxicity: relationship to risk assessment*, Environmental Health Perspectives, 102 (1994), pp.157-167.

[BC] R. Bellman and K.L. Cooke, "Differential-Difference Equations," Rand Corporation, 1963.

[BGS] H.T. Banks, D.S. Gilliam, and V.L. Shubov, *Global solvability for damped abstract nonlinear hyperbolic systems*, Differential and Integral Equations, 10 (1997), pp. 309-332.

[BIW] H.T. Banks, K. Ito, and Y. Wang, *Well-posedness for damped second order systems with unbounded input operators*, CRSC-TR93-10, June 1993, Differential and Integral Equations, 8 (1995), pp. 587-606.

[BM] H.T. Banks and C.J. Musante, *Well-posedness for a class of abstract nonlinear parabolic equations with time delay*, in preparation.

[BMT] H.T. Banks, C.J. Musante, and H.T. Tran, *A dispersion model for the hepatic uptake and elimination of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)*, in preparation.

[BSW] H.T. Banks, R.C. Smith, and Y. Wang, "Smart Material Structures: Modeling Estimation, and Control," J. Wiley & Sons, 1996.

[NS] A.W. Naylor and G.R. Sell, "Linear Operator Theory in Engineering and Science," Springer-Verlag, New York, 1982.

[PG] A. Poland and E. Glover, *Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol*, Journal of Biology and Biological Chemistry, 251 (1976), pp. 4936-4946.

[PTG] A. Poland, P. Teitelbaum, and E. Glover, *[^{125}I]2-Iodo-3,7,8-trichlorodibenzo-p-dioxin-binding species in mouse liver induced by agonists for the Ah receptor: characterization and identification*, Molecular Pharmacology, 36 (1989), pp. 113-120.

[RR1] M.S. Roberts and M. Rowland, *A dispersion model of hepatic elimination: 1. formulation of the model and bolus considerations*, Journal of Pharmacokinetics and Biopharmacokinetics, 14 (1986), pp. 227-260.

[SDG] S.A. Skene, I.C. Dewhurst, and M. Greenberg, *Polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans: the risks to human health, a review*, Human Toxicology, 8 (1989), pp. 173-203.

[V] M. Van den Berg, J. De Jongh, H. Poiger, and J. Olson, *The toxicokinetics and metabolism of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) and their relevance for toxicity*, Critical Reviews in Toxicology, 24 (1994), pp. 1-74.

[VA] R. Voorman and S.D. Aust, *Specific binding of polyhalogenated aromatic hydrocarbon inducers of cytochrome P-450d to the cytochrome and inhibition of its estradiol-2-hydroxylase activity*, Toxicology and Applied Pharmacology, 90 (1987), pp. 69-78.

[Wl] J. Wloka, "Partial Differential Equations," Cambridge University Press, 1992.

Investigation of the Iron-Bearing Phases of the Columbus AFB Aquifer

David C. Powell
Graduate Student
Physical Sciences Department

College of William and Mary
School of Marine Science
Gloucester Point, Virginia 23062

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory
Tyndall AFB, Florida

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1997

Investigation of the Iron-Bearing Phases of the Columbus AFB Aquifer

David C. Powell
Graduate Student
Physical Sciences Department
College of William and Mary/SMS

Abstract

This work was a part of the USAF Natural Attenuation Study (NATS) to determine whether natural attenuation is a cost effective and practical remediation technique for organic contaminants in aquifers. It was part of a multifaceted experiment designed to investigate the availability of the iron-bearing phases for use by indigenous, iron-reducing microbes in the degradation of organic groundwater contaminants at Columbus Air Force Base (CAFB), Columbus, Mississippi. The research focused on three topics: (1) the characterization of the iron phases in the CAFB aquifer materials by selective extraction methods; (2) method development for working with anaerobes, for future laboratory studies; and (3) demonstration of the abiotic production of magnetite, thus indicating that magnetite is not an exclusive indicator of microbial degradation under iron-reducing conditions.

Investigation of the Iron-Bearing Phases of the Columbus AFB Aquifer

David C. Powell

Introduction

Natural attenuation, defined as any naturally-occurring process that decreases the concentration of contaminants in groundwater, has been proposed as viable, low cost, technique for remediation of hydrocarbons in contaminated aquifers. For natural attenuation to be considered feasible, biodegradation must be shown to be taking place and the maximum extent of the dissolved contaminant plume migration must be estimated (Borden et al., 1997). The Natural Attenuation Study (NATS), conducted by the United States Air Force (USAF) Armstrong Laboratory, Environmental Quality Laboratory (AL/EQL), is designed to examine the effects of several naturally-occurring processes, including microbial degradation, dilution, and evaporation, on the transport and fate of an organic contaminant plume. As the goal of remediation design is to reduce the contaminant concentration in the aquifer to a level that no longer threatens human health or the environment (Nyer, 1997), natural attenuation may be the best choice remediation design in areas where other remediation techniques will cause significant environmental damage (Boggs et al., 1993).

Uncontrolled release of hydrocarbons contaminating subsurface systems is widespread. Leaking underground storage tanks, improper disposal of hazardous materials in landfills and excessive use of agricultural chemicals are all important sources of contamination (Mihelcic et al., 1995). In most cases, contaminants dissolved in the water column are of primary importance. NATS focuses on contamination by petroleum derived fuels, especially the primary dissolved components known as BTEX (benzene, toluene, ethylbenzene and xylene). BTEX makes up less than 20 percent of fuel mixtures like JP-4 jet fuel, gasoline and diesel fuel, but accounts for 82 percent and 98 percent of the dissolved compounds from fresh JP-4 and gasoline, respectively (Wiedemeier, et al., 1996).

Unlike other previously known iron-reducing bacteria that only partially breakdown hydrocarbons through a fermentation pathway, *Geobacter metallireducens*, is the first

microorganism known to completely, anaerobically reduce aromatic hydrocarbons to CO_2 using Fe (III) as a terminal electron receptor (Lovley and Phillips, 1988). It also is the first pure culture capable of anaerobically-reducing toluene, a common component of petroleum products such as gasoline and jet fuel (Lovley and Phillips, 1988; Lovley and Lonergan, 1990). More recently, other bacteria have been cultured that use Fe (III) to completely oxidize hydrocarbons. For example, *Shewanella alga* strain BrY grows in a medium with synthetic goethite as the only oxidant, reducing the Fe (III) by 8-18% (Roden and Zachara, 1996) and *Pelobacter carbinolicus* is able to use both Fe (III) and S^0 as terminal oxidants (Lovley et al., 1995).

The most common evidence of intrinsic bioremediation includes: (1) documented loss of contaminants at the field site, (2) chemical and geochemical data, including increased concentration of intermediates and metabolites of anaerobic metabolism of BTEX, and concomitant decreased concentration of terminal electron acceptors in the aquifer, and (3) laboratory culture studies, showing microbial degradation (Wiedemeier et al., 1996; Wilson and Madsen, 1996). In addition, there are several other important parameters that can be used as evidence for bioremediation. For example, differences between the $\delta^{13}C$ values of hydrocarbons and indigenous carbon sources (e.g. plant matter, soil carbonates) can be exploited to trace the origins of metabolic end products (Conrad et al., 1997). Also, the concentration of dissolved H_2 has been shown to correspond to the reduction-oxidation (redox) potential and is used to determine which terminal oxidation processes predominate at specific locations in an aquifer (Lovley et al., 1994).

Magnetite formation may be another possible indicator of microbial degradation. Magnetite has been observed in contaminated aquifers under iron reducing conditions and is considered to be a by-product of microbial iron reduction. Iron-reducing bacteria use ferric iron as a terminal electron acceptor, with an overall reaction of reducing solid Fe (III) oxides to aqueous Fe (II). Magnetite is considered to be formed by a matrix rearrangement and water expulsion when aqueous Fe (II) sorbs into amorphous or poorly crystalline Fe (III) oxides on the surface of the aquifer material. If magnetite is formed only in the presence of iron-reducing microbes, then magnetite production may be an indicator of microbial degradation of hydrocarbons. This study was designed to test for abiotic production of magnetite under similar culture conditions required for microbial production of magnetite.

There is concern that highly crystalline iron (III) oxides may not be available for microbial remediation of hydrocarbons. Some studies of aquatic sediments and submerged soil show that crystalline iron (III) oxides like goethite (α-FeOOH) and hematite (α-Fe_2O_3) may not be reducible by microorganisms, but amorphous and poorly crystalline iron (III) oxides are readily reduced. (Phillips et al., 1993; Lovley and Phillips, 1987) Other studies have shown that Fe (III)-reducing bacteria may be able to survive and produce significant quantities of Fe (II) in anaerobic soils and subsurface environments where crystalline iron (III) oxides such as goethite are the dominant forms of Fe (III) (Roden and Zachara, 1996). Difference in utilization may be a function of the naturally-occurring consortium of bacteria. Regardless of their source mineralogy, reduction of even small amounts of iron (III) oxides can cause significant changes in the geochemistry of an aquifer. Reduction may release soluble Fe (II) and other adsorbed species (e.g. phosphates and trace metals) or provide a reactive Fe (II) surface capable of participating in secondary redox or mineral-forming reactions (Roden and Zachara, 1996).

This study investigated changes in the iron-bearing phases in the Columbus AFB aquifer. The study site has low concentrations of sulfate, nitrate and manganese, so microbially-mediated iron (III) reduction of the emplaced hydrocarbons will be the primary oxidant once the oxygen has been depleted. Specifically, this study characterizes changes in the Fe (III)-bearing phases due to microbial degradation, assesses the availability of these phases to support degradation processes, and examine the use of magnetite as a potential indicator of microbial degradation.

Field Site Description

The NATS experiment started with the emplacement of mixed hydrocarbon coated onto 30 m^3 of cleaned sand at approximately 15% saturation on December 1 1995. The aquifer site (Figure 1) has been sampled for sediments four times since the emplacement: January, April and September 1996, and March 1997. The January 1996 samples provide the background data to which all other samples will be compared[1]. A distinct, migrating hydrocarbon plume has developed with the highest concentrations about the middle depth. Samples analyzed from this

[1] This is justified because the hydrocarbon plume had not migrated significantly down gradient in the 40 day period from the emplacement to the sampling date in January 1996.

depth are the most likely to show changes in the iron-bearing phases because of their enhanced contact with the hydrocarbon plume; most subsequent samples analyzed were from this depth. Select samples from other areas from the aquifer were used as a baseline in conjunction with the January 1996 samples.

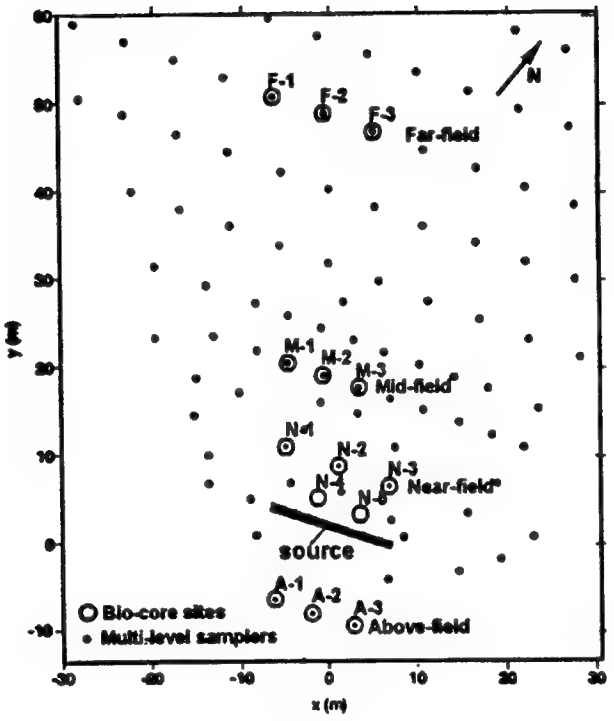


Figure 1: Aquifer site at CAFB

Methodology

Iron Extraction

Ferrous and ferric iron were extracted from the sediments in order to determine the amount of these iron phases present in the aquifer samples. Amorphous iron was extracted using 1 N HCl (Chou and Zhou, 1983). Specifically, 0.25 g sediment and 25 ml N2-sparged 1N HCl were mixed in an anaerobic glovebox. The sediment and acid were sealed in serum vials with Teflon coated septum and placed in a 70°C water bath for 30 minutes. The supernatant was then

sampled by syringe, to reduce exposure to oxygen, and filtered through a 0.45 μm syringe filter.

Total Fe (III) oxides were extracted by reductive dissolution using citrate-dithionite-bicarbonate (CDB) (Heron et al., 1994). For this, 0.25 g sediment, 20 ml 0.3 M Na citrate and 2 ml 1.0 M $NaHCO_3$ were mixed in a centrifuge tube and heated in a water bath to 80°C; before adding 0.5g $Na_2S_2O_4$. The sample was kept at 80°C for at least 15 minutes, with periodic mixing. The samples are centrifuged at 7000rpm for 10 minutes and the supernatant decanted for analysis. This procedure was repeated three times to obtain quantitative results. An alternative method for extracting total Fe (III) oxides was also evaluated. This procedure consisted of a single extraction of 0.25 g sediment with 20 ml of 5 N HCl for two weeks at room temperature (Heron et al., 1994). Each extraction was replicated twice. The efficiency of each method was evaluated by powder x-ray diffraction (XRD) analysis and flame atomic adsorption (FAA) analysis; the method that reproducibly yielded the most extractable iron will be used for all future extractions.

The supernatants from the selective extraction procedures were analyzed to determine the amount of iron released. Fe (II) was determined by complexing 0.5 ml of extract supernatant with 5 ml of 0.004 M ferrozine buffered with HEPES and analyzing on a UV/Vis spectrophotometer at 562 nm. Fe (II) was only analyzed from the 1 N HCl extractions, because the CDB extraction process reduces Fe (III) to Fe (II) and because 5 N HCl would have overwhelmed the buffer. Total iron was determined from the HCl and CDB extractions by FAA.

Anaerobic Cultures

Future experiments are being designed to degrade acetate and toluene on the iron-phases present on homogenized sediments from the CAFB aquifer system by a mono-culture, *Geobacter metallireducens,* and with indigenous microbes from CAFB. Procedures, previously used by collaborators, for working with strict anaerobes in laboratory experiments were tested for use in these future experiments. The procedure consisted of combining 100 ml of an acetate medium, 10 ml each of mineral and vitamin supplements and 2 ml of an iron gel[2] into 150 ml serum vials. The mixture was purged for 30 minutes using an O_2-free N_2/CO_2 (70/30%) gas mixture. The vials

[2] Prepared by dissolving 54 g $FeCl_3$ in 500 ml H_2O and adjusting the pH to 7 with 10 M NaOH. The precipitate was washed and centrifuged 6 times to remove the Na^+ and Cl^- ions from the gel.

were then sealed with thick rubber stoppers, crimped shut and autoclaved at 125°C and 15 lb. pressure for 20 minutes. Once cooled, a 1 ml inoculate of *G. metallireducens* was added using sterile transfer techniques[3]. The vials were then stored at 30 °C.

Magnetite Crystallization

For the magnetite formation experiments, batch cultures of *G. metallireducens* were grown under strict anaerobic condition with and without the addition of 50 to 800 mg/L of Fe (II). Abiotic controls were incubated along with the bacterial cultures. The cultures and abiotic controls were prepared in serum vials using the culturing method described above. After autoclaving and cooling, the vials were doped with autoclaved, anaerobic solutions of Fe (II) to provided final concentrations of 50 to 800 mg/L Fe (II). Vials containing ≥400mg/L Fe (II) began visibly releasing gas bubbles within 24 hours. Thus, all the vials were transferred to an anaerobic glovebox and vented. Parallel experiments were conducted to determine the source and identity of the gas. In these experiments, 5 ml of Fe gel and ~800ppm Fe (II) were mixed with different constituents: buffered water, acetate medium, or acetate medium with the vitamin and mineral supplements. The generated gasses were analyzed by gas chromatography using a size exclusion, packed column and a thermo-couple detector.

Results & Discussion

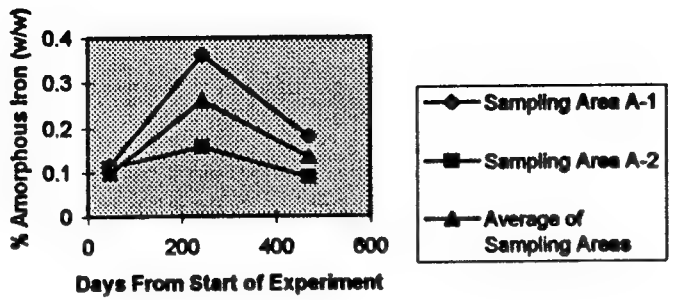
From the total iron extractions completed to date, it is clear that the CDB and 5 N HCl extractions yield different results. The CDB extraction yielded iron concentrations that varied by as much as 53% for sample duplicates and appeared to alter the aquifer material, forming clay phases not found in the original aquifer material. Fe analysis variability for duplicate samples from the HCl extractions typically was <5% and did not exceed 28%. Additional 5 N HCl extractions will be performed to determine if the total iron changes in the aquifer over the duration of the NATS experiment. The amorphous extractions show changes in the iron-phases in the near field over time, but not to a degree that allows separation of the changes from the natural heterogeneity

[3] The rubber stoppers are cleaned with methanol and the residue flamed off. Sterile needles and syringes are flushed with a O_2-free N_2/CO_2 gas mixture run through a sterilized glass syringe packed with sterilized glass wool.

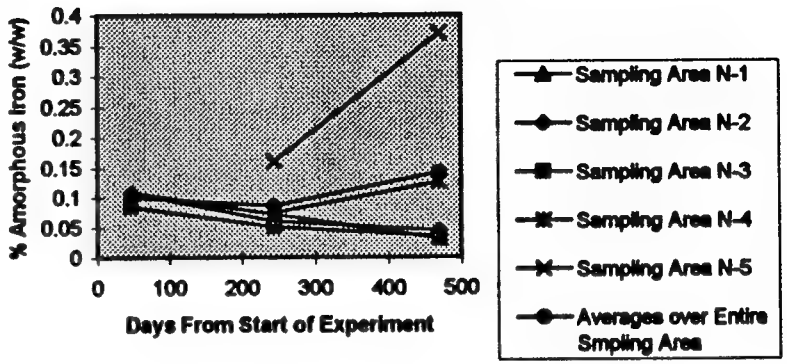
of the system (Figure 2). Perhaps, combining the amorphous data along with the total iron data yet to be gathered as well as additional data from the September 1997 sampling currently being conducted will allow a better understanding of the use of the iron-phases by the naturally-occurring bacteria in the CAFB site.

G. metallireducens is a well characterized vigorous monoculture that readily reduces amorphous iron to Fe(II); one distinguishing characterization of its growth is formation of magnetite in vitro (Lovely and Lonergan, 1995). Similarly, this study produced healthy cultures as evidenced by rapid Fe(II) production and formation of magnetite in the vials. In contrast, abiotic controls did not produce Fe(II) or magnetite.

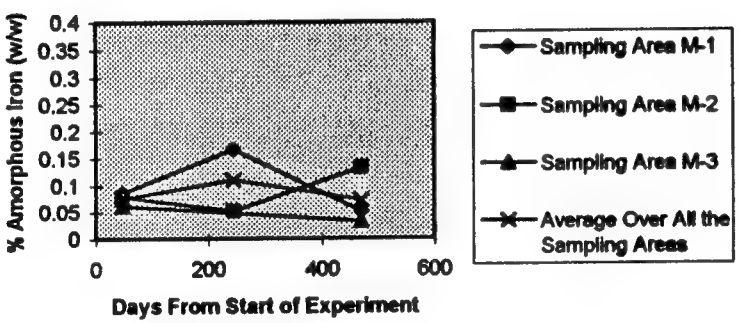
To assess if magnetite is formed only in the presence of iron-reducing microbes, *G. metallireducens* cultures grown in the presence of excess Fe(II) were compared with abiotic controls. Magnetite formed in all vials with the microbes and in abiotic vials containing at least 200 mg/L Fe(II). However, the production of magnetite was much faster and more extensive in the presence of the *G. metallireducens.* H_2 gas was evolved during these experiments. The hydrogen is likely produced due to incorporation of Fe(II) and the rearrangement of the iron gel to form magnetite. Formation of magnetite in the abiotic microcosms demonstrates that magnetite crystallization is a chemical process occurring independent of microbial activity. While magnetite production in itself may not be a sufficient indicator of microbial degradation of hydrocarbons, under appropriate aquifer conditions where there is no other source of Fe(II), it might be combined with other indicators to show anaerobic degradation of hydrocarbons.



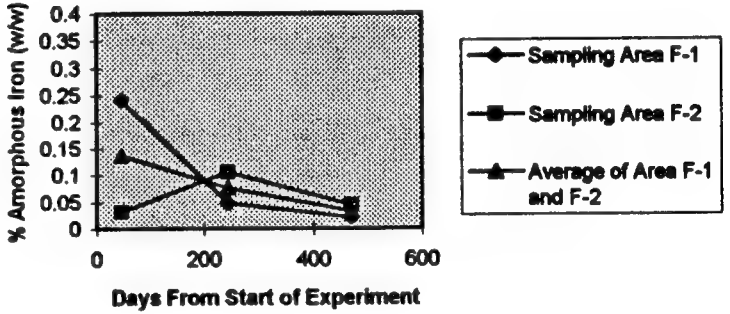
A



B



C



D

Figure 2: Weight Percent of Amorphous Iron Middle Levels (59.5-58.5 m above sea level):

A) Above Field; B) Near Field; C) Middle Field; D) Far Field

References

Boggs, J. M., L. M. Beard, W. R. Waldrop, T. B. Stauffer, W. G. MacIntyre, C. P. Antworth Transport of Tritium and Four Organic Compounds During a Natural-Gradient Experiment (MADE-2), *EPRI Topical Report* TR-101998, Electric Power Research Institute, Palo Alto, Ca.

Borden, R. C., R. A. Daniel, L. E. LeBrun IV, and C. W. Davis. 1997. Intrinsic biodegradation of MTBE and BTEX in a gasoline-contaminated aquifer. *Water Resources Research* 33(5):1105-1115.

Chao, T.T. and L. Zhou, . 1984. Extraction techniques for selective dissolution of amorphous iron oxides from soils and sediments. *Soil Science Society of America Journal* 47:225-232.

Conrad, M. E., P. F. Daley, M. L. Fischer, B. B. Buchanan, T. Leighton, and M.Kashgarian. 1997. Combined ^{14}C and $\delta^{13}C$ Monitoring of *in Situ* Biodegradation of Petroleum Hydrocarbons. *Environmental Science and Technology* 31:1463-1469.

Heron, G., C. Crouzel, A. C. M. Bourg, and T.H. Christensen. 1994. Speciation of Fe(II) and Fe(III) in contaminated aquifer sediments using chemical extraction techniques. *Environmental Science and Technology* 26:1698-1705.

Lovley, D.R., E. J. P. Phillips, D. J. Lonergan, and P. K. Widman. 1995. Fe(III) and S^0 Reduction by *Pelobacter carbinolicus. Applied and Environmental Microbiology.* 61(6):2132-2138.

Lovley, D. R., and D. J. Lonergan. 1990. Anaerobic Oxidation of Toluene, Phenol, and *p*-Cresol by the Dissimilatory Iron-Reducing Organism, GS-15. *Applied and Environmental Microbiology.* 56(6):1858-1864.

Lovley, D. R., and E. J. P. Phillips, 1988. Novel Mode of Microbial Energy Metabolism: Organic Carbon Oxidation Coupled to Dissimilatory Reduction of Iron or Manganese. *Applied and Environmental Microbiology.* 54(6):1472-1480.

Lovley, D. R., F. H. Chapelle, and J. C. Woodward, 1994. Use of Dissolved H_2 Concentrations to Determine Distribution of Microbially Catalyzed Redox Reaction in Anoxic Groundwater. *Environmental Science and Technology* 28(7):1205-1210.

Mihelcic, J. R., A. Pritschow and D. R. Leuking. 1995. Uptake of Dissolved and Oil Phase Organic Chemicals by Bacteria. *Groundwater Monitoring Review* Summer:100-106.

Murphy, E. M., T. R. Ginn, A. Chilakapati, C. T. Resch, J. L. Phillips, T. W. Wietsma, and C. M. Spadoni. 1997. The Influence of Physical Heterogeneity on Microbial Degradation and Distribution in Porous Media. *Water Resources Research* 33(5):1087-1103.

Nyer, E. K., and M. J. Gearhart. 1997. Plumes don't Move. *Groundwater Monitoring Review* Winter:100-106.

Phillips, E. J. P., D. R. Lovley, D. J. Lonergan, and E. E. Roden. 1993. Composition of Non-Microbially Reducible Fe(III) in Aquatic Sediments. *Applied and Environmental Microbiology*. 59(8):2727-2729.

Reinhard M., S. Shang, P. K. Kitanidis, E. Orwin, G. D. Hopkins and C. A. Lebron. 1997. In Situ BTEX Biotransformation under Enhanced Nitrate- and Sulfate-Reducing Conditions. *Environmental Science and Technology* 31:28-36.

Roden, E. E. and J. M. Zachara. 1996. Microbial Reduction of Crystalline Iron(III) Oxides: Influence of Oxide Surface Area and Potential for Cell Growth. *Environmental Science and Technology* 30:1618-1628.

Wiedemeier, T. H., M. A. Swanson, J. T. Wilson, D. H. Kampbell, R. N. Miller, and J. E. Hansen. 1996. Approximation of Biodegradation Rate constants for Monoaromatic Hydrocarbons (BTEX) in Ground Water. *Groundwater Monitoring Review* Summer:186-194.

Wilson, M. S. and E. L. Madsen. 1996. Field Extraction of transient Intermediary Metabolite indicative of real Time *in Situ* Naphthalene Biodegradation. *Environmental Science and Technology* 30:2099-2103.

THE EFFECT OF VISUAL SIMILARITY AND REFERENCE FRAME ALIGNMENT ON THE RECOGNITION OF MILITARY AIRCRAFT

Christopher Schreiner
Graduate Student
Department of Psychology

Miami University
Oxford, OH. 45056

Final Report for:
Graduate Student Research Program
Brooks AFB
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Armstrong Laboratory

August 1997

THE EFFECT OF VISUAL SIMILARITY AND REFERENCE FRAME ALIGNMENT ON THE RECOGNITION OF MILITARY AIRCRAFT

Mr Christopher Schreiner
Department of Psychology
Miami University
Oxford Ohio 45056 USA

ABSTRACT

Aircraft similar in appearance (homogeneous) and dissimilar in appearance (heterogeneous) were studied at orientations consistent with the environmental frame of reference (canonical) or inconsistent with the environmental frame of reference (non-canonical). Response time data for correct identifications indicate that identification performance was better for heterogeneous than homogeneous aircraft. This performance advantage for heterogeneous aircraft was found at both the original training orientations and for novel orientations. Canonical orientations during learning produced better identification performance than non-canonical orientations. Implications for aircraft recognition training are discussed.

THE EFFECT OF VISUAL SIMILARITY AND REFERENCE FRAME ALIGNMENT ON THE RECOGNITION OF MILITARY AIRCRAFT

Mr Christopher Schreiner

INTRODUCTION

When teaching aircraft recognition, what cognitive processes can be utilized to maximize recognition effectiveness? In this paper, we address two such processes. First, we consider the overall visual similarity of the set of aircraft to be taught. Second, we consider the alignment of the internal reference frame of the aircraft with the environmental reference frame.

The general context in which we study these processes is that of object constancy across orientation manipulations within the picture plane. We measure recognition performance as a function of angular disparity between aircraft images presented during learning and later at test. This manipulation provides a measure of learning transfer to novel orientations. This is a crucial issue in aircraft training because aircraft are typically viewed in many more orientations than other stimulus classes. An optimal training regimen for the identification of aircraft would incorporate the least number of views necessary to achieve high object constancy across novel orientations. The number of views, and the specific views chosen, may vary depending upon the overall visual similarity of the group of aircraft to be learned, as well as the initial orientations used during training.

Visual similarity can be modeled as category membership. That is, the more dissimilar a group of objects, the "higher" level category they compose, whereas objects more similar in appearance can be considered "lower" level category members. For example, working with aircraft, a novice may classify a wide variety of aircraft as simply being "planes," whereas an Air Force pilot may identify aircraft by their specific names and even model number (F15-Eagle, or more specifically, F15c-Eagle).

Rosch, Mervis, Gray, Johnson, & Boyes-Braem (1976) suggested that expertise may interact with the basic level of categorization that people use to identify classes of objects. Since their suggestion, a growing body of evidence suggests that expertise promotes facility with category levels subordinate to the basic level. For example, Palmer, Jones, Hennessy, Unze, and Pick (1989) found that musicians and nonmusicians differed on the level at which they identified musical instruments. The musicians' identification responses were more specific than the nonmusicians ("violin" as compared to "strings," or "trumpet" as compared to "brass"). Tanaka and Taylor (1991) found that experts on birds and dogs were as fast and accurate at making subordinate level identifications (sparrow, beagle) as they were at the basic level (bird, dog). In comparison, novices on birds and dogs showed a decrement in identification performance at the subordinate level as compared to the basic level. Tanaka and Taylor (1991) concluded that increased expertise led to increased use of, and access to, subordinate level information. (See also Johnson and Mervis (1997), and Gauthier and Tarr (1997))

In our experiment, we used subjects who by self report were novices at recognizing aircraft. Thus past research suggested that their recognition for aircraft that were similar in appearance (lower category) should be

slower and less accurate than that for objects that were dissimilar in appearance (higher category). We predicted that as subjects gained expertise with the aircraft, the disparity between their identification performance on similar and dissimilar aircraft would lessen, and perhaps disappear completely.

Aircraft have a distinct axis of symmetry that is also the major axis of elongation. There are data that recognition performance is enhanced when the internal reference frame of the object is aligned with the environmental reference frame. For example, Shiffrar and Shepard (1991) presented two successive cubes on a CRT and asked subjects to determine if the rotations of the two cubes in three-dimensional space were the same or different. Both the speed and accuracy of the subjects' performance was best when the axis of symmetry of the cube was aligned with the vertical axis of the environmental reference frame. Similarly, Pani (1993) instructed subjects to predict the outcomes of various rotations of a square in three-dimensional space. Subjects' accuracy was best when the rotations where normal to both the square and the environmental reference frame (See also Appelle (1972); Corballis, Nagourney, Shetzer, & Stefanatos (1978); Friedman and Hall (1996); Hinton (1988); Parsons (1995); Pani (1994); Pani and Dupree (1994); Pani, Jeffres, Shippey, & Schwartz (1996); Pani, Williams, & Shippey (1995)).

We tested this finding with aircraft images. We predicted that when the major axis of the aircraft were aligned with the 0, 90, 180, and 270 degree orientations of the environmental reference frame, subsequent recognition performance would be better than when these reference frames were not aligned. We also tested the transfer of recognition performance as a function of object and environmental reference frame alignment. Would improved recognition performance due to aligned object and environmental frames lead to improved transfer across orientations?

Whitmore, P. G., Rankin, W. C., Baldwin, R. O., & Garcia, S. (1972) collected data relevant to the issue of learning transfer to novel orientations. They attempted to determine the minimum number of aircraft training views necessary to achieve an identification accuracy rate of 95% for novel orientations. Using slides made from scaled models, they trained subjects using regularly sequenced orientations in three-dimensional space, and then collected accuracy data for identification of slides made from models at novel orientations. Their results suggested that as few as nine well chosen views could facilitate accurate identification at 45 novel views. They reported anecdotal evidence suggesting that subjects took much longer to make identification responses when the angular disparity between novel views and test views was large.

We contend that due to the potentially severe consequences of inaccurate aircraft identification it is prudent to collect response time data on transfer of learning to novel orientations. Given enough time, accurate identification responses can be made based upon very few learning views. However, the nature of aircraft recognition precludes lengthy analysis of the target image. Thus, with this study, we focus on response time for correct responses as the primary dependent measure.

METHODS

Subjects

82 subjects were recruited through several local temporary agencies using the criteria that they 1) were between 18 and 30 years of age, 2) had at least a high school diploma or Graduate Equivalency Degree (GED), and 3) reported no previous experience with recognizing aircraft. When they arrived at the laboratory all subjects were

screened for visual acuity using a standard Schnelling Eye Chart. Only those subjects whose vision was 20/30 or better participated. All subjects were paid for their work.

Materials

Stimuli consisted of 12 high-resolution, gray-scale images of military aircraft. A "0 degree" orientation was designated for each aircraft and was a top view with the nose of the aircraft being oriented towards the top of the CRT. Thus the aircraft appeared to be standing on its tail and pointing straight up. Sixty-four images of each of the 12 aircraft were made by rotating the 0 degree image in the picture plane at increments of 5.625 degrees. When presented on the CRT in the 0 degree orientation, the height of the aircraft was approximately 10 cm. Because subjects sat approximately 60 cm from the CRT, the stimulus images subtended a visual angle of approximately 9.5 degrees.

A mask image was also generated. The mask was a circular image with a gradient fill from black to white, starting with black in the center and becoming progressively lighter along the radius, thus giving the effect of a dark centered starburst from black to gray to white across concentric circles.

Based upon the results of cluster analyses from previous experiments, and cluster analyses from pilot studies, four aircraft that were similar (Homogeneous) in appearance were chosen and four aircraft that were dissimilar (Heterogeneous) in appearance were chosen. The four Homogeneous aircraft were from the same cluster (see Figure 1), whereas the four Heterogeneous aircraft were from different clusters (see Figure 2). These eight aircraft were studied by the subjects for later identification. An additional four aircraft were chosen as distractors. Two distractors were chosen for the Homogeneous set based upon the experimenters' subjective determination that they resembled the aircraft in the Homogeneous set. Two distractors were chosen for the Heterogeneous set based upon the experimenters' subjective determination that they were as distinct as the aircraft in the Heterogeneous set. Thus there were six Homogeneous aircraft (four targets and 2 distractors) and six Heterogeneous aircraft (four targets and 2 distractors). As noted above, each of the twelve aircraft could be presented at any of 64 orientations in the picture plane.

The actual military names of the Homogeneous aircraft were Eagle, Hornet, Fulcrum, and Flanker as target stimuli with Foxbat and Falcon as the distracters. The actual military names of the Heterogeneous aircraft were Fishbed, Galab, Tomcat, and Farmer as target stimuli with Mirage and Fresco as the distracters.

A canonically-oriented set of aircraft consisted of all eight of the target aircraft described above oriented at 0, 90, 180, and 270 degrees in the picture plane. See Figure 1 for examples of aircraft at the four canonical orientations. A non-canonically-oriented set of aircraft consisted of all eight of the target aircraft described above oriented at 22.5, 112.5, 202.5, and 292.5 degrees in the picture plane. See Figure 2 for examples of aircraft at the four non-canonical orientations.

The stimuli were presented on an IBM compatible microcomputer with a high-resolution RGB CRT. Responses were made on the keyboard. Accuracy and millisecond response times were recorded by the computer. Each participant was tested in a three-sided cubicle.

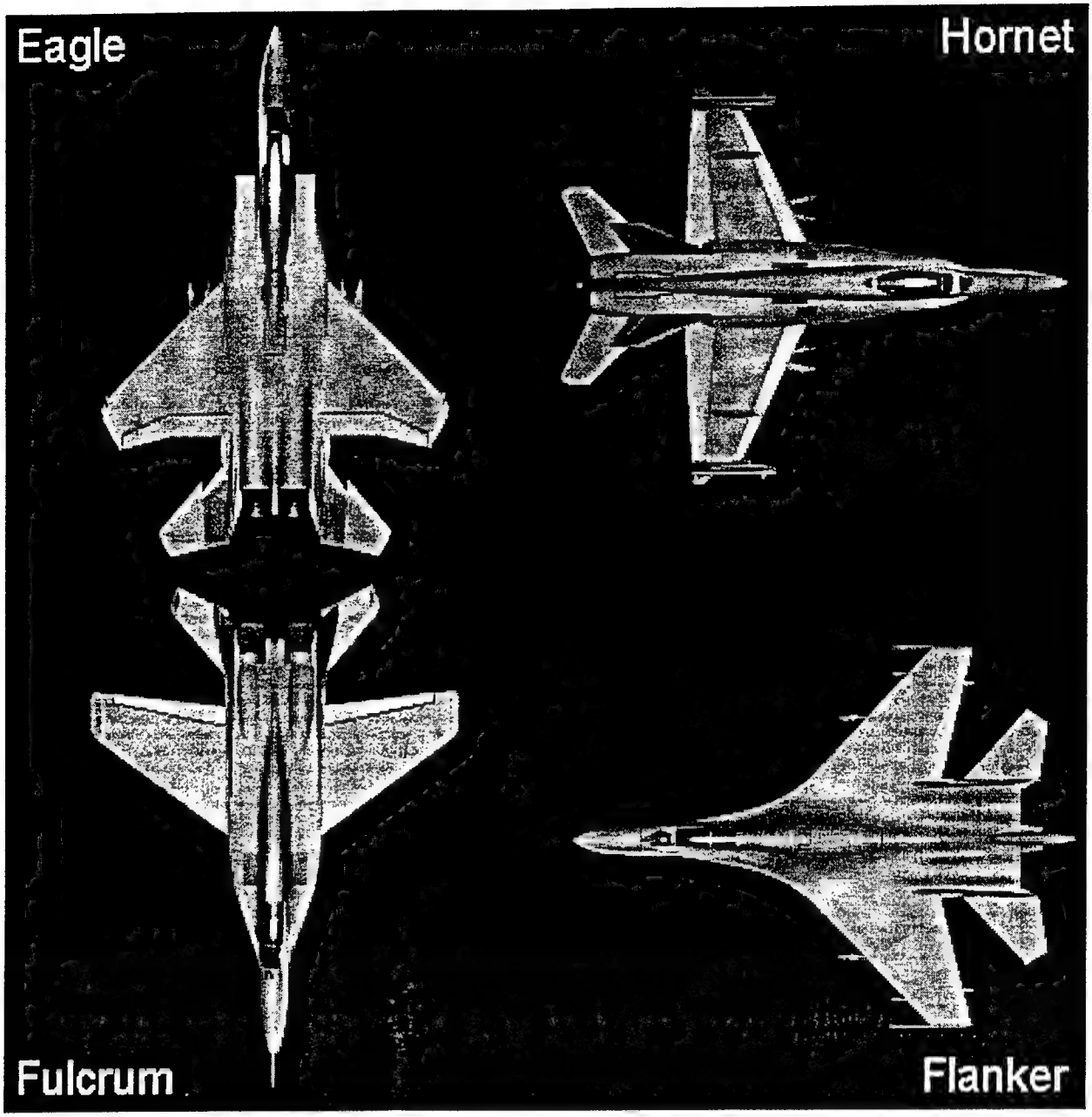


Figure 1. The four aircraft that were homogeneous in appearance, at the four canonical learning orientations, shown with their names.

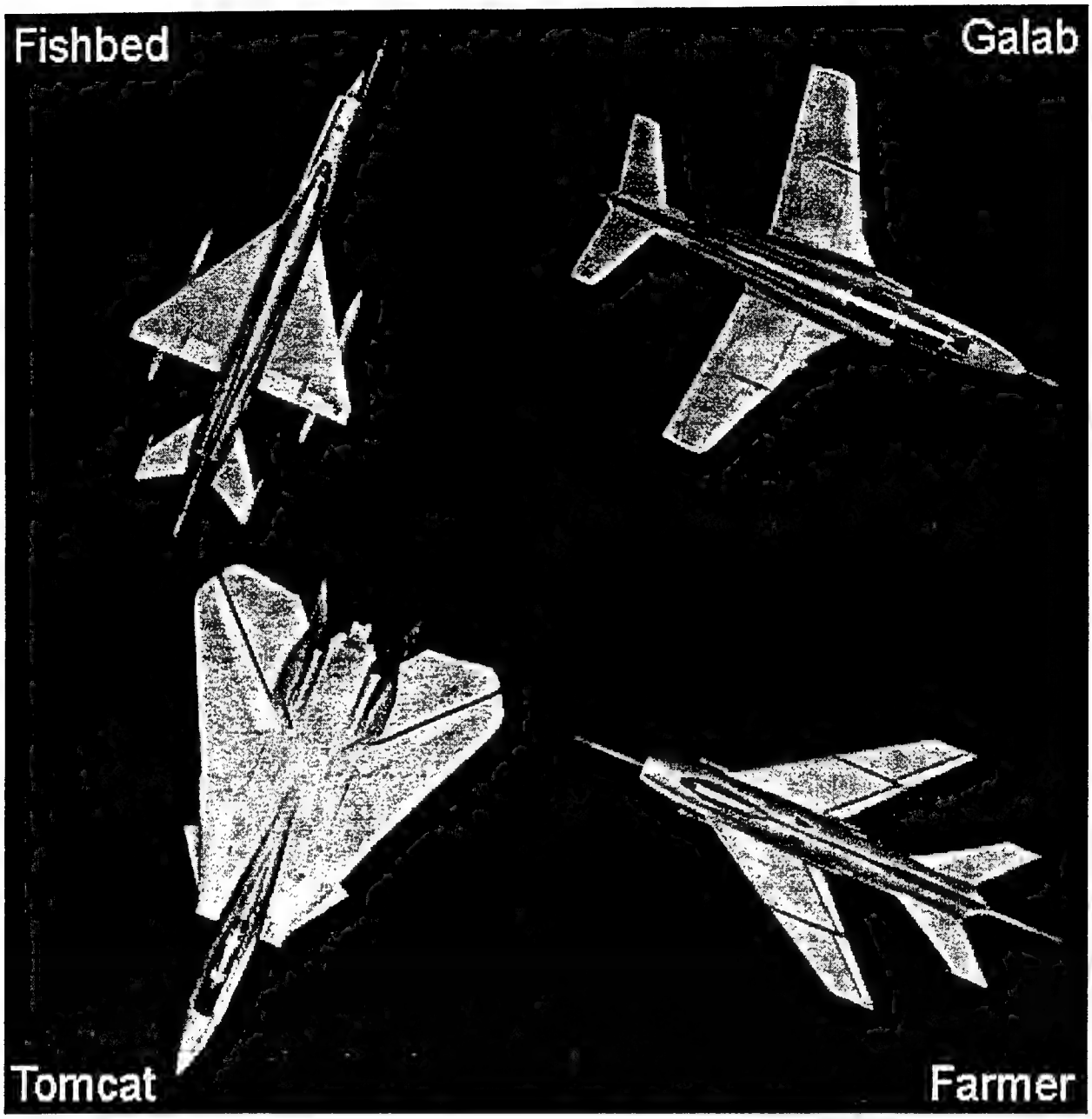


Figure 2. The four aircraft that were heterogeneous in appearance, at the four non-canonical learning orientations, shown with their names.

Design and Procedure

Pretest/Posttest. Immediately before and after the experiment all subjects completed a multiple-choice aircraft identification test. This test contained 19 aircraft including the 12 aircraft used as stimuli. Three images of an individual aircraft were presented on the CRT along with seven possible answers. Six of the answers were names of aircraft, while the seventh answer was always "none of the above." Responses were made on the keyboard using the number keys 1 - 6 for the aircraft names, and 7 for "none of the above." After a response was made, another aircraft was presented, and so on, until all 19 aircraft had been presented.

There were two forms of the test. In Form A the three images depicted the standard three views used by the Department of Defense to teach aircraft recognition. These were a side view (side), a top view (top), and a head-on view (front). In Form B all three views were oblique. In one view the aircraft was pointing towards the viewer with the nose angled up and to the right (up), in another the aircraft was pointing towards the viewer with the nose angled down and to the right (down), and in the third the aircraft was pointing away from the viewer with the nose angled up and to the right (away).

All subjects completed both forms of the test before the experiment, and again after the experiment. The order of the two versions was counterbalanced across subjects with an ABBA/BAAB design.

Aircraft Naming The experiment had two factors, each with two levels, resulting in four conditions. The first factor (Learning Orientation) varied the orientation at which the aircraft stimuli were learned. The two levels of the Learning Orientation factor were Canonical (0, 90, 180, and 270 degrees) and Non-canonical (22.5, 112.5, 202.5, and 292.5 degrees). Learning Orientation was a between-subjects factor. The second factor (Similarity) varied the degree of visual similarity among aircraft to be learned. The two levels of the Similarity factor were Homogeneous and Heterogeneous. Similarity was a within-subjects factor, and was counterbalanced across subjects. Thus the four experimental conditions were 1) Canonical and Homogeneous, 2) Canonical and Heterogeneous, 3) Non-canonical and Homogeneous, and 4) Non-canonical and Heterogeneous.

The procedure was identical for each of the four conditions. There was a Learning Phase followed by a Testing Phase. Each phase was composed of blocks of trials. Trials consisted of an aircraft stimulus both preceded and followed by a 100 ms mask. The duration of the aircraft stimulus was 5000 ms or until a response was made, whichever came first. The responses were made on keyboard keys that were labeled with the aircraft names. There was a fifth key labeled "NA" for none of the above.

Learning Phase The Learning Phase consisted of nine blocks of trials. During the first block the four aircraft were presented ten times at each of the four learning orientations. Thus there were 160 trials in the first block (4 aircraft X 4 orientations X 10 repetitions). The order of trials was randomized per subject within the block. On each trial the name of the aircraft was presented along with the aircraft image. Subjects were instructed to study the image for the full 5000 ms, and to associate the name with the aircraft. After the second mask cleared the CRT, the subjects pressed the key labeled with the correct aircraft name. Subjects were instructed to concentrate and strive for perfect accuracy because their recognition of the four planes would be tested later. Accuracy feedback was provided by a beep if an incorrect key was pressed.

The second, third, and fourth blocks of the Learning Phase were identical. During these blocks the four aircraft were presented ten times at each of the four learning orientations. Thus there were 160 trials in each block (4 aircraft X 4 orientations X 10 repetitions). The order of trials was randomized per subject per block. On each trial only the aircraft image was presented. The aircraft names were not presented. Subjects were instructed to press the key labeled with the correct aircraft name as soon as they recognized each aircraft. Subjects were instructed to respond as quickly as possible while still maintaining high accuracy. Accuracy feedback was provided by a beep if an incorrect key was pressed. After the fourth block subjects were given a ten-minute rest break.

The fifth block was identical to the first block except for the number of repetitions. The four aircraft were presented once at each of the four learning orientations. Thus there were 16 trials in the fifth block (4 aircraft X 4 orientations). The order of trials was randomized per subject within the block. On each trial the name of the aircraft was presented along with the aircraft image. Subjects were instructed to study the image for the full 5000 ms, and to associate the name with the aircraft. After the second mask cleared the CRT, the subjects pressed the key labeled with the correct aircraft name. Subjects were instructed to concentrate and strive for perfect accuracy because their recognition of the four planes would be tested later. Accuracy feedback was provided by a beep if an incorrect key was pressed.

The sixth, seventh, eighth, and ninth blocks were identical to the second, third, and fourth blocks. During these blocks the four aircraft were presented ten times at each of the four learning orientations. Thus there were 160 trials in each block (4 aircraft X 4 orientations X 10 repetitions). The order of trials was randomized per subject per block. On each trial only the aircraft image was presented. The aircraft names were not presented. Subjects were instructed to press the key labeled with the correct aircraft name as soon as they recognized each aircraft. Subjects were instructed to respond as quickly as possible while still maintaining high accuracy. Accuracy feedback was provided by a beep if an incorrect key was pressed. After the ninth block subjects were given a ten minute rest break.

Test Phase The Test Phase consisted of two blocks of trials. The first block was identical to the fifth block of the Learning Phase. The four aircraft were presented once at each of the four learning orientations. Thus there were 16 trials in the first block (4 aircraft X 4 orientations). The order of trials was randomized per subject within the block. On each trial the name of the aircraft was presented along with the aircraft image. Subjects were instructed to study the image for the full 5000 ms, and to associate the name with the aircraft. After the second mask cleared the CRT, the subjects pressed the key labeled with the correct aircraft name. Subjects were instructed to concentrate and strive for perfect accuracy because their recognition of the four planes would be tested later. Accuracy feedback was provided by a beep if an incorrect key was pressed. The order of trials was randomized per subject within the block.

The second block consisted of 320 trials during which the primary data for the experiment were collected. During the second test block the four aircraft were presented once at each of the 64 test orientations for a total of 256 presentations (4 aircraft X 64 orientations). In addition, two distractor aircraft were presented. One distractor was presented once at each of the 32 odd-numbered orientations (1, 3, 5, etc.), while the other distractor was presented once at each of the 32 even-numbered orientations (2, 4, 6, etc.), for a total of 64 distractor presentations (2 distractors X 32 orientations). The order of trials was randomized per subject within the block. On each trial only the

aircraft image was presented. The aircraft names were not presented. Subjects were instructed to press the key labeled with the correct aircraft name as soon as they recognized each aircraft. Subjects were informed that distractors had been added, and they were to press the key labeled "NA" for none of the above at the presentation of a distractor. Subjects were instructed to respond as quickly as possible while still maintaining high accuracy. Unlike all proceeding blocks in the Learning Phase and the Test Phase, no accuracy feedback was provided if an incorrect key was pressed.

The second block was preceded by 24 practice trials for the purpose of stabilizing the subjects' response time performance before the introduction of the primary experimental trials. Subjects were not informed that the first 24 trials of the second block were practice trials. For the practice trials the four aircraft were presented once at each of the four learning orientations for a total of 16 presentations (4 aircraft X 4 orientations). In addition, the two distractors were presented once at each of the 45, 135, 225, and 315 degree orientations for a total of eight presentations (2 distractors X 4 orientations). The order of the practice trials was randomized per subject.

RESULTS

All response times are expressed in milliseconds. For correct responses, medians were calculated across the four target aircraft at each orientation, for each block, for each subject. The seven blocks of the Learning Phase were analyzed separately from the final block of the Test Phase.

Learning Phase

Because there were four orientations in each of the seven learning blocks, and each subject participated in both levels of the Similarity factor, there were 56 medians per subject in the learning phase. The Similarity factor had the two levels of Homogeneous and Heterogeneous and was a within-subjects factor. The Block factor had seven levels (there were seven learning blocks) and was a within-subjects factor. The Orientation factor had four levels (there were four orientations used during learning) and was a within-subjects factor. The Learning Orientation factor had the two levels of Canonical and Non-canonical and was a between-levels factor. The design for the response time ANOVA for the Learning Phase was: ((2) Learning Orientation) X ((2) Similarity) X ((7) Block) X ((4)Orientation).

The ANOVA revealed significant main effects for Learning Orientation, $F(1, 80) = 4.54$, $p < .05$, $\underline{MSE} = 545751.40$; for Similarity, $F(1, 80) = 86.83$, $p < .001$, $\underline{MSE} = 110663.30.23$; for Block, $F(6, 480) = 120.80$, $p < .001$, $\underline{MSE} = 14128.42$; and for Orientation, $F(3, 240) = 5.78$, $p < .01$, $\underline{MSE} = 3063.52$.

As shown in Figure 3, for the Learning Orientation factor, the response time for Non-canonical presented aircraft ($\underline{M} = 830$, $\underline{SE} = 15.42$) was higher than the response time for Canonically presented aircraft ($\underline{M} = 784$, $\underline{SE} = 15.42$).

As shown in Figure 4, for the Similarity factor, the response time for Homogeneous aircraft ($\underline{M} = 852$, $\underline{SE} = 13.74$) was higher than the response time for Heterogeneous aircraft ($\underline{M} = 761$, $\underline{SE} = 9.84$).

As shown in both Figures 3 and 4, for the Block factor, the response times decreased monotonically until the fourth block, after which they held steady through the remainder of the learning phase($\underline{M1} = 905$, $\underline{SE1} = 15.41$, $\underline{M2} = 840$, $\underline{SE2} = 12.90$, $\underline{M3} = 814$, $\underline{SE3} = 11.61$, $\underline{M4} = 762$, $\underline{SE4} = 10.40$, $\underline{M5} = 770$, $\underline{SE5} = 10.47$, $\underline{M6} = 781$, $\underline{SE6}$

= 10.08, M7 = 777, SE7 = 10.15.

For the Orientation factor, the four average response times are not meaningful because they are collapsed across the Canonical and Non-canonical orientations. Specifically, the first Orientation factor represents a collapse of performance at the 0 and 22.5 degree orientations, the second is that of the 90 and 112.5 degree orientations, the third is that of the 180 and 202.5 degree orientations, while the fourth is that of the 270 and 292.5 degree orientations. However, all terms that include the Orientation X Learning Orientation factor are meaningful because this interaction does not collapse across the Canonical and Non-Canonical factor.

The ANOVA revealed significant interactions for Similarity X Block, $F(6, 480) = 27.89$, $p < .001$, MSE = 9736.86, and Similarity X Block X Orientation, $F(18, 1440) = 1.98$, $p < .01$, MSE = 1566.85. However, the latter interaction cannot be interpreted due to the manner in which the Orientation factor is collapsed. There were no other significant interactions.

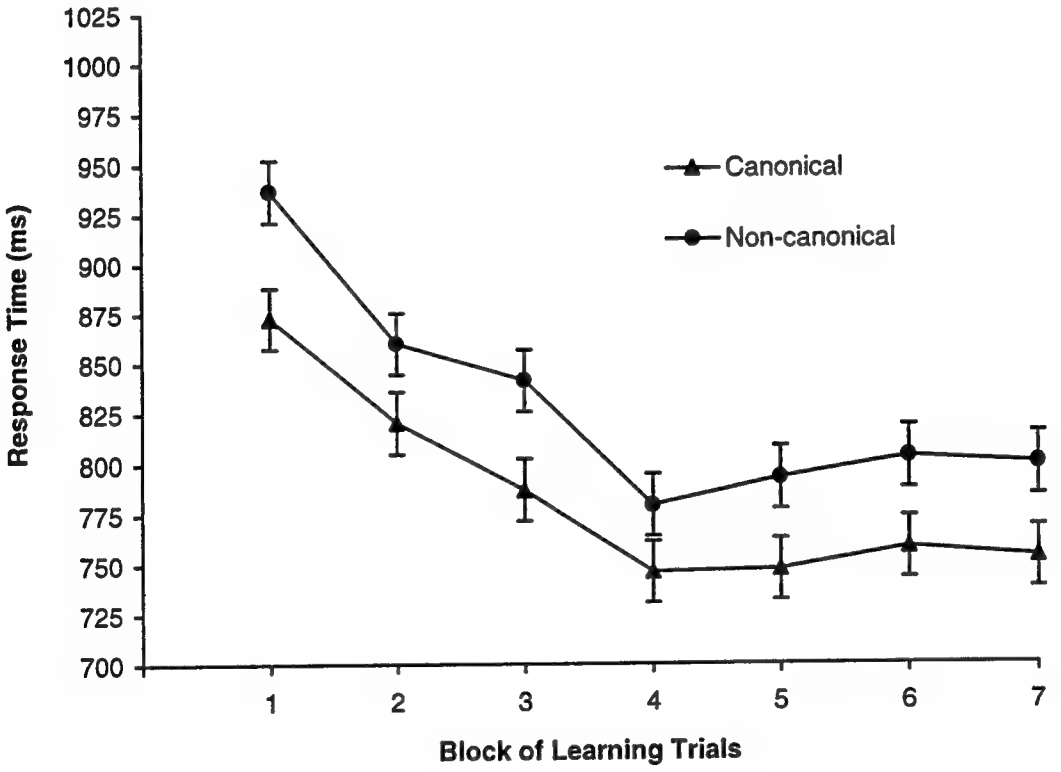


Figure 3. Response time for correct recognition of aircraft images during learning as a function of learning block. Canonical and Non-canonical learning orientations plotted separately. Error bars indicate the normalized standard error.

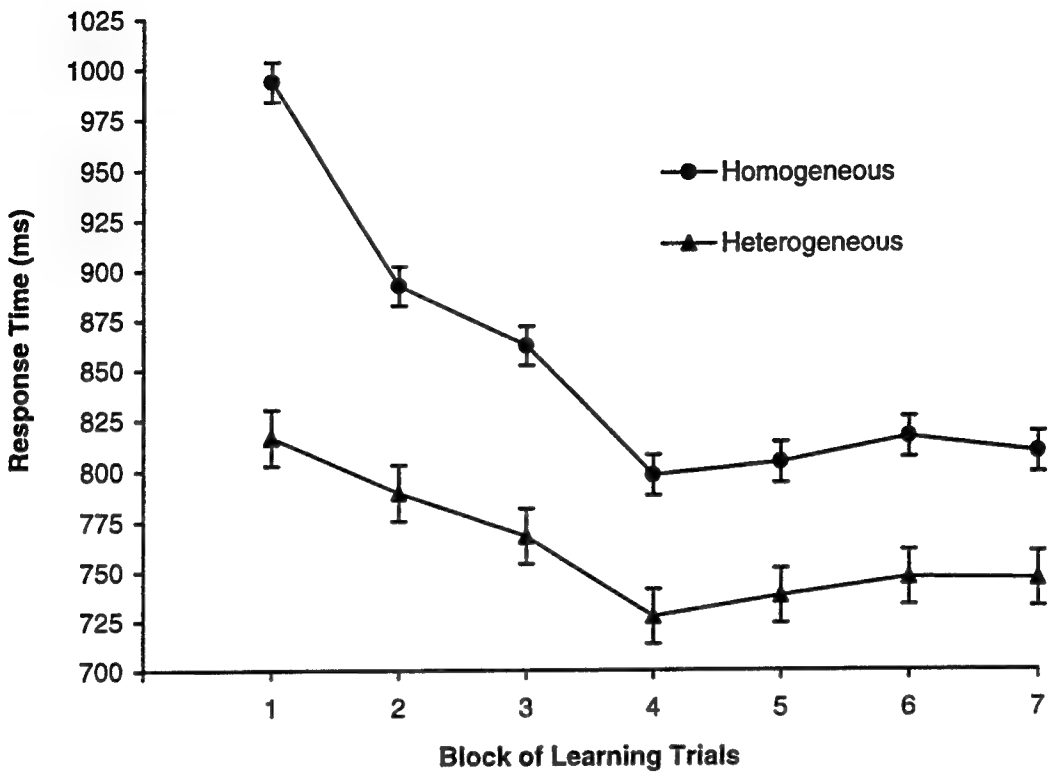


Figure 4. Response time for correct recognition of aircraft images during learning as a function of learning block. Homogeneous and Heterogeneous aircraft plotted separately. Error bars indicate the normalized standard error.

Test Phase

Because there were 64 orientations at which the aircraft were presented during the Test Phase, there were 64 medians per subject in the test phase. The 64 medians were collapsed as a function of distance from the learning orientations. Thus, all medians that were 5.625 degrees from a learning orientation were averaged, all medians that were 11.250 degrees from a learning orientation were averaged, all medians that were 16.875 degrees from a learning orientation were averaged, and so forth. Because the learning orientations were separated by 90 degrees, the maximum distance that the collapsed data could be from a learning orientation was 45 degrees. After collapsing, the data represented median response times per subject at 0.000, 5.625, 11.250, 16.875, 22.500, 28.125, 33.750, 39.375, and 45.000 degrees from a learning orientation, and made up the nine levels of the within-subjects Orientation factor. The Learning Orientation factor had the two levels of Canonical and Non-canonical and was a between-levels factor. The Similarity factor had the two levels of Homogeneous and Heterogeneous and was a within-subjects factor. The design for the response time ANOVA for the Test Phase was: ((2) Learning Orientation) X ((2) Similarity) X ((9)Orientation).

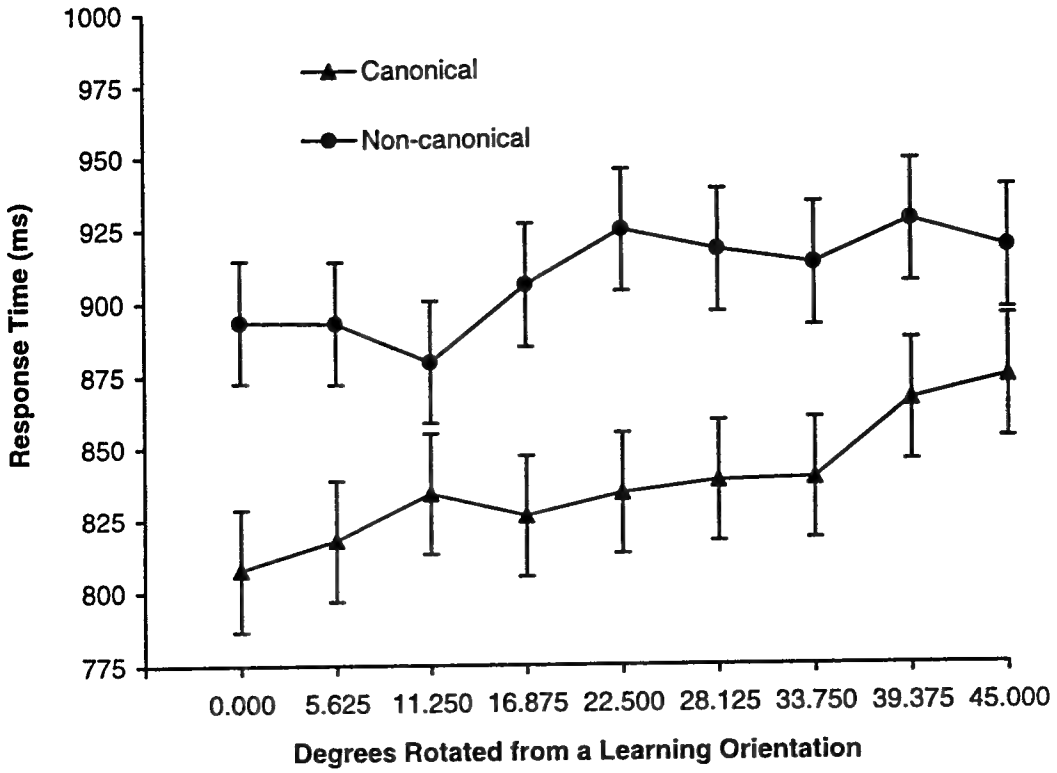


Figure 5. Response time for correct recognition of aircraft images as a function of degree rotated in the picture plane away from a learning orientation. Canonical and Non-canonical learning orientations plotted separately. Error bars indicate the normalized standard error.

The ANOVA revealed significant main effects for Learning Orientation, $F(1, 80) = 5.65$, $p < .05$, $MSE = 322875.60$; for Similarity, $F(1, 80) = 57.87$, $p < .001$, $MSE = 60913.23$; and for Orientation, $F(8, 640) = 9.88$, $p < .001$, $MSE = 4639.13$.

As shown in Figure 5, for the Learning Orientation factor, the response time for Non-canonical presented aircraft ($M = 907$, $SE = 20.92$) was higher than the response time for Canonically presented aircraft ($M = 837$, $SE = 20.92$).

As shown in Figure 6, for the Similarity factor, the response time for Homogeneous aircraft ($M = 921$, $SE = 17.82$) was higher than the response time for Heterogeneous aircraft ($M = 823$, $SE = 14.22$).

As shown in both Figures 5 and 6, for the Orientation factor, the response time tended to increase monotonically as the distance from the learning orientation increased ($M1 = 850$, $SE1 = 14.45$, $M2 = 855$, $SE2 = 14.07$, $M3 = 857$, $SE3 = 13.90$, $M4 = 866$, $SE4 = 15.77$, $M5 = 879$, $SE5 = 15.14$, $M6 = 877$, $SE6 = 15.07$, $M7 = 875$, $SE7 = 15.17$, $M8 = 896$, $SE8 = 17.14$, $M9 = 895$, $SE9 = 19.14$).

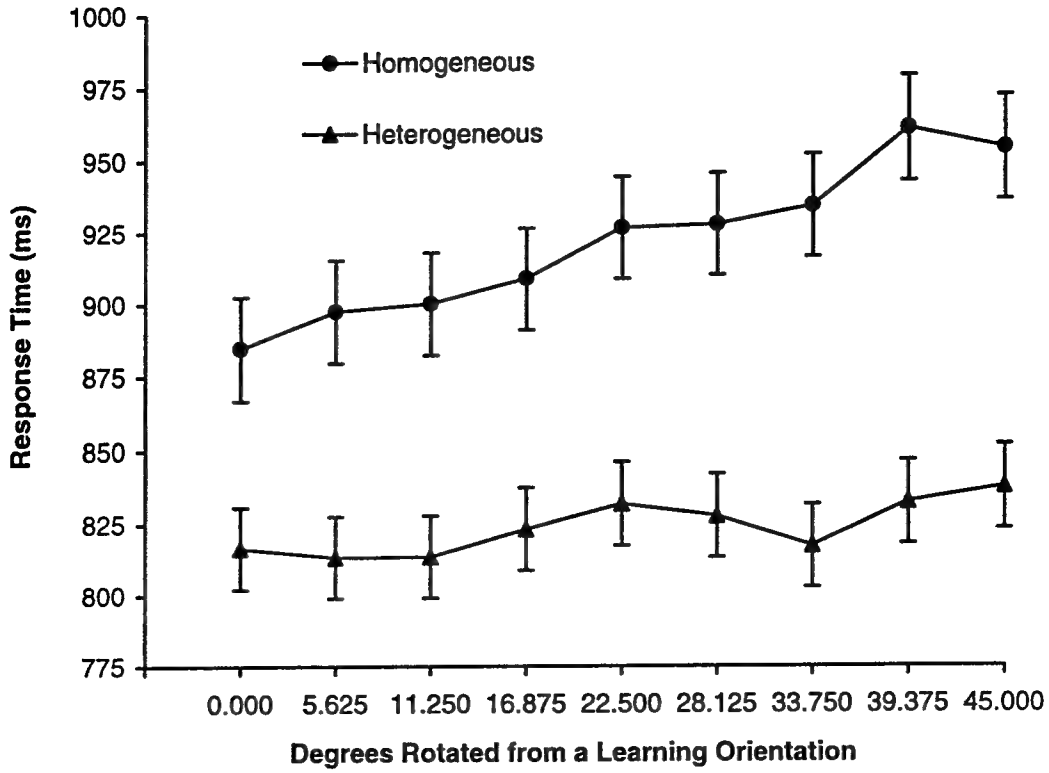


Figure 6. Response time for correct recognition of aircraft images as a function of degree rotated in the picture plane away from a learning orientation. Homogeneous and Heterogeneous aircraft plotted separately. Error bars indicate the normalized standard error.

The ANOVA revealed significant interactions for Learning Orientation X Orientation, $F(8, 640) = 2.47$, $p < .05$, $MSE = 4639.13$, and Similarity X Orientation, $F(8, 640) = 3.06$, $p < .01$, $MSE = 5026.07$. There was no Learning Orientation X Similarity interaction, $F(1, 80) < 1.00$.

Polynomial contrasts on the Orientation factor revealed that there was a significant linear trend, $F(1, 80) = 43.27$, $p < .001$, $MSE = 7812.31$, with response time increasing as distance from a learning orientation increased. The sixth-order polynomial contrast was significant but could not be interpreted.

The linear component of the Orientation factor was tested between the Canonical and Non-canonical levels of the Learning Orientation factor. There was no significant linear X Learning Orientation interaction.

The linear component of the Orientation factor was tested between the Homogeneous and Heterogeneous levels of the Similarity factor. There was a significant linear X Similarity interaction, $F(1, 80) = 13.26$, $p < .001$, $MSE = 8279.96$.

DISCUSSION

Interpretation of the results is straightforward, and the results from the Test Phase mirror those of the Learning Phase.

Discrimination between aircraft that are dissimilar in appearance is easier than between aircraft that are similar. During the Learning Phase, the Homogeneous aircraft were consistently slower to be identified than the Heterogeneous aircraft. Even after performance had leveled off around the fourth learning block, the Heterogeneous aircraft maintained a consistent advantage of approximately 65-70 ms. This difference was slightly greater at test, approximately 100 ms.

A similar pattern was found for the Learning Orientation manipulation. During the Learning Phase, the Non-canonically presented aircraft were consistently slower to be identified than the Canonically presented aircraft. Even after performance had leveled off around the fourth learning block, the Canonically presented aircraft maintained a consisted advantage of approximately 40-45 ms. This difference was slightly greater at test, approximately 70 ms.

The Orientation factor measured the transfer of recognition from the aircraft at the learning orientations to adjacent orientations in the picture plane. There was a linear decrease in recognition performance as the angular disparity increased between learning orientation and test orientation.

The average visual similarity of the set of aircraft interacted with the transfer of recognition to novel orientations. For the Heterogeneous aircraft, there was very little performance degradation as the angular disparity increased between learning orientation and test orientation. The Homogeneous aircraft generated considerably more performance degradation than the Heterogeneous aircraft for the same measure.

In contrast, there was no interaction between Learning Orientation and transfer of recognition to novel orientations. The Canonically and Non-canonically presented aircraft generated the same amount of transfer of learning.

We conclude that both average similarity of the set of aircraft and the initial orientation at which they are learned, have a powerful effect on identification performance both during learning and later at test.

The data suggest that the more distinctive the individual aircraft are within a set, the more transfer of learning will occur from the initial learning orientation to novel orientations. This in turn suggests that for distinctive aircraft fewer total views of aircraft would be needed during learning to insure accurate identification at novel orientations. In comparison, aircraft that are more similar in appearance will require more total views to insure accurate identification at novel views.

Further research is required to quantify the relationship between homogeneity of the aircraft stimulus set and the minimum number of view required to maximize recognition performance across all novel orientations. Based upon our current research (Ashworth, 1997), a mixture of oblique and accidental views (looking down any foreshortened axis) would be required. Because we have noted that the learning transfers very poorly between accidental views, all accidental views would need to be included in any training regimen, regardless of homogeneity. Thus based upon the results reported here, a savings in training time and resources could be attained by using fewer oblique views during learning of heterogeneous aircraft as compared to homogeneous aircraft.

This prescription mirrors that of Gibson (1947). In his work on aircraft recognition for the Army Air Corp, Gibson (1947) suggested that aircraft of similar appearance should be shown in pairs while students made timed discrimination responses between the pairs. He further stated that this procedure was not necessary for aircraft that were dissimilar in appearance. Thus Gibson also suggested different teaching methodologies based upon visual similarity of the aircraft.

Considering past research, (e.g., Appelle (1972); Corballis, Nagourney, Shetzer, & Stefanatos (1978); Friedman and Hall (1996); Hinton (1988); Parsons (1995); Pani (1993); Pani (1994); Pani and Dupree (1994); Pani, Jeffres, Shippey, & Schwartz (1996); Pani, Williams, & Shippey (1995); Shiffrar and Shepard (1991)) the main effect for initial learning orientation is not surprising. However, our data reveal just how powerful this effect is. For example, After four blocks of the seven-block Learning Phase, identification performance was asymptotic for both the Canonically-oriented and Non-canonically-oriented aircraft. This suggests that continued practice would not lead to equivalent performance across this manipulation.

Further, inspection of test data in Figure 3 suggests that identification performance at 22.5 degrees from a Canonically-oriented training orientation is better than that at 0.0 degrees from a Non-canonically-oriented training orientation. This means that training at the Canonical orientations and then testing at the Non-canonical orientations produces better identification than training at the Non-canonical orientations and then testing at the Non-canonical orientations! This surprising result is significant, $F(1, 80) = 4.05$, $p < .05$, $\underline{MSE} = 35868.26$.

The mechanism which drives the enhanced performance on aircraft studied at the canonical orientations is unknown. For example in his review of the reference frame literature, Palmer (1989) concludes that object recognition is mediated primarily by the intrinsic reference frame for objects whose structure clearly defines one, by an environmental reference frame for objects whose structure does not clearly define an intrinsic reference frame, with the best recognition occurring when intrinsic reference frames align with the environmental reference frame (see also Palmer, Rosch, & Chase (1981)). However Palmer does not elaborate upon why the coincidence of these two reference frames produces the best recognition performance. He simply notes that it is assumed that "there are biases toward picking other salient orientations as the reference orientation, especially gravitational vertical or the top-bottom axis of the retina" (p.127).

We concur with Shephard (1982, 1984) and Shiffrar and Shephard (1991), that this effect must ultimately derive from the invariant nature of the terrestrial gravitational field and its predictable influence on objects. Objects with primary internal axes are the most stable if their axes are either in line with gravity, or orthogonal to gravity. Objects whose primary internal axes are at oblique angles to gravity tend to topple over. Our experience with objects in the world is predominated by objects whose internal axes are normal to gravity, and thus our perceptual apparatus exploits these regularities in the world.

The results suggest that if only a very limited number of views are to used during training then views should be chosen such that the intrinsic and environmental reference frames align. Further, if the external constraints of the training regimen will allow the optimum number of views to be used, then those in which the intrinsic and environmental reference frames align should be included. As already noted, accidental views should always be included.

REFERENCES

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The oblique effect in man and animals. *Psychological Bulletin, 78,* 266-278.

Ashworth, A. R. S., III (1997). A comparison of oblique and accidental views for the learning of military aircraft. Manuscript in preparation.

Corballis, M. C., Nagourney, B. A., Shetzer, L. I., & Stefanatos, G. (1978). Mental rotation under head tilt: Factors influencing the location of the subjective reference frame. Perception & Psychophysics, 24(3), 263-273.

Friedman, A., & Hall, D. L. (1996). The importance of being upright: Use of environmental and viewer-centered reference frames in shape discriminations of novel three-dimensional objects. Memory & Cognition, 24(3), 285-295.

Gauthier, I., & Tarr, M. J. (in press). Becoming a greeble expert: Exploring mechanisms for face recognition. *Vision Research.*

Gibson, James, J., & Gagne, Robert, M. (1947). Research on the recognition of aircraft. In James J. Gibson (Ed.), Motion picture testing and research (pp. 113-168). *Army Air Forces Aviation Psychology Program Research Reports, Report No. 7.* U.S. Government Printing Office, Washington 25, D.C.

Hinton, G. E., & Parsons, L. M. (1988). Scene-based and viewer-centered representations for comparing shapes. *Cognition, 30,* 1-35.

Johnson, K. E., Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126(3),* 248-277.

Palmer, S. E. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp, S. Ballesteros (Eds.), *Object Perception: Structure and Process* (pp. 121-163). Hillsdale, NJ: Erlbaum.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long, A. Baddeley (Eds.), *Attention & Performance IX* (pp. 135-151). Hillsdale, NJ: Erlbaum.

Palmer, C. F., Jones, R. K., Hennessy, B. L., Unze, M. G., & Pick, A. D. (1989). How is a trumpet known? The "basic object level" concept and the perception of musical instruments. *American Journal of Psychology, 102,* 17-37.

Pani, J. R. (1993). Limits on the comprehension of rotational motion: mental imagery of rotations with oblique components. *Perception, 22,* 785-808.

Pani, J. R. (1994). The generalized cone in human spatial organization. *Spatial Vision, 8(4),* 491-501.

Pani, J. R., & Dupree, D. (1994). Spatial reference systems in the comprehension of rotational motion. *Perception, 23,* 929-946.

Pani, J. R., William, C. T., & Shippey, G. T. (1995). Determinants of the perception of rotational motion: Orientation of the motion to the object and to the environment. *Journal of Experimental Psychology: Human Perception and Performance, 21(6),* 1441-1456.

Pani, J. R., Jeffres, J. A., Shippley, G. T., & Schwartz, K. J. (1996). Imagining projective transformations: Aligned orientations in spatial organization. *Cognitive Psychology, 31,* 125-167.

Parsons, L. M. (1995). Inability to reason about an object's orientation using an axis and angle of rotation. *Journal of Experimental Psychology: Human Perception and Performance, 21(6),* 1259-1277.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Shepard, R. N. (1982). Perceptual and analogical bases of cognition. In J. Mehler, E. C. T. Walker, & M. Garrett (Eds.), *Perspectives on mental representation* (pp. 49-67). Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1984). Ecological constraints o internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91,* 417-447.

Shiffrar, M. M., & Shepard, R. N. (1991). Comparison of cube rotations around axes inclined relative to the environment or to the cube. *Journal of Experimental Psychology: Human Perception and Performance, 17(1),* 44-54.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457-482.

Whitmore, P. G., Rankin, W. C., Baldwin, R. O., & Garcia, S. (1972). Studies of aircraft recognition training. *Human Resources Research Organization Technical Report 72-5,* HumRRO Division 5, Fort Bliss, Texas.

OH RADICAL REACTION RATE CONSTANT AND PRODUCT STUDY OF 2-PROPOXYETHANOL

John N. Sempeles
Masters Candidate
Environmental Engineering Sciences Department

University of Florida
PO Box 116450
Gainesville, FL 32611

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

OH RADICAL REACTION RATE CONSTANT AND PRODUCT STUDY OF 2-PROPOXYETHANOL

John N. Sempeles
Masters Candidate
Environmental Engineering Sciences Department
University of Florida

Abstract

The rate constant for the reaction of 2-propoxyethanol and OH radicals was studied using the relative rate technique. A preliminary study of the products was also made. 2-Propoxyethanol was placed in a Teflon bag with a reference compound, methyl nitrate, nitric oxide, and compressed air. The mixture was photolyzed and subsequently analyzed via GC - FID. Four reference compounds (n-nonane, n-hexane, dodecane, and 1-heptanol) were used to improve the accuracy of the rate constant. The runs done with n-nonane and n-hexane were so erratic that they were discarded. The two references that were used gave little overlap even considering statistical error in 95% confidence interval. Runs with dodecane received a rate constant of $27.2 \pm 4.4 \times 10^{-12}$ $cm^3 molecule^{-1} s^{-1}$ and runs with 1-heptanol received a rate constant of $19.0 \pm 3.8 \times 10^{-12}$ $cm^3 molecule^{-1} s^{-1}$. Experimental results indicated that a systematic error was occurring each day but was random in between days. Several changes were made to the equipment to narrow down the cause of the problem, however the problem still appeared. Future suggestions are to replace the collection trap, where active sites might be forming and causing compounds to stick, and perhaps try a different column.

OH RADICAL REACTION RATE CONSTANT AND PRODUCT STUDY OF 2-PROPOXYETHANOL

John N. Sempeles

Introduction

The study of oxygenated volatile organic compounds such as 2-propoxyethanol has many beneficial purposes. The Air Force uses glycol ethers similar to 2-propoxyethanol for paint stripping and re-coating operations. These compounds, when released to the air, react to form tropospheric ozone as well as aldehydes, ketones, nitrates, and other products which might be considered environmental hazards. Since reaction with OH radicals is a major loss mechanism for 2-propoxyethanol, rate constant studies give a good indication of the expected tropospheric lifetime of the compound. Product studies suggest possible mechanisms for reactant loss and product yields. Characterizing these ethers' reactions in the air serves the purposes of determining how the Air Force can best meet environmental regulations, both locally and globally.

Methodology

The reaction of 2-propoxyethanol with OH radicals was conducted in a Teflon® film chamber. Calibrated volumes of reactants were introduced into an approximately 100 L Teflon® bag using a cajon T with a septum. Dry, compressed air was added as a diluent. The Teflon bag was placed in a chamber surrounded by the following mix of lamps: 6 - Philips TL40W/D3; 1 - GE F40BL; 2 - QPANEL UV351; and 7 - QPANEL UV340.

The wavelength range of light emitted was 300 - 450 nm. Samples were taken out of the bag using a 6.4 mm Swagelock fitting and measured using a 0 - 100 L $min.^{-1}$ mass flow controller. The kinetic experiments were carried out at room temperature (298 K) and at atmospheric pressure (760 torr).

All samples were analyzed using a Hewlett Packard (HP) 5890 gas chromotograph (GC) with a flame ionization detector (FID). Gas samples for GC analysis were cryogenically collected on a Hastelloy C sample trap (ca. 1.3 ml) and injected onto the GC column RESTEK RTX - 200 column (0.53 mm, 30 m, 3.0 μm film thickness) with a 10 port heated rotary valve.[1] The GC temperature program used was initially: 30 °C for 3 min. and 14 °C/min. to 150°C for 5 min. This program was later changed to 35 °C for 5 min. and 10 °C/min. to 220 °C for 2 min. for the majority of the experiments (see results section). The GC column pressure program utilized was: 2 psi. for 500 minutes. Helium, UHP grade from Air Products Inc., was used as the carrier gas.

Hydroxyl radicals were generated from the photolysis of methyl nitrite in air containing NO :

$$CH_3ONO + h\nu \rightarrow CH_3O + NO \quad (1)$$

$$CH_3O + O_2 \rightarrow HCHO + HO_2 \quad (2)$$

$$HO_2 + NO \rightarrow OH + NO_2 \quad (3)$$

CH_3ONO was prepared in gram quantities using the method of Taylor et al. [2] and stored in a lecture bottle at room temperature. The CH_3ONO purity was verified by GC/MS/FTIR.

Rate coefficients of OH reaction with 2 - propoxyethanol

The rate coefficient for the gas phase reaction of 2-propoxyethanol (TST compound) with OH reduced was determined using the relative rate method.

$$TST + OH \rightarrow Products \quad (4)$$

$$REF + OH \rightarrow Products \quad (5)$$

Assuming that reaction with OH radicals is the only loss mechanism, the rate equations for these two reactions can be integrated and combined:

$$\ln [(TST)_0 / (TST)_t] = (k_{TST} / k_{REF}) \ln [(REF)_0 / (REF)_t] \quad (6)$$

where the subscripts 0 and t refer to concentrations taken at time zero and time t, respectively. A plot of $\ln [(TST)_0 / (TST)_t]$ vs. $\ln [(REF)_0 / (REF)_t]$ gives the ratio of rate constants k_{TST} / k_{REF} as the slope. Multiplying this slope value by the reference's rate constant gives 2-propoxyethanol's rate constant. Four reference compounds, n-nonane, n-hexane, dodecane and 1-heptanol, were used in this study to increase the accuracy of the 2-propoxyethanol / OH rate constant.

The concentrations of reactants in the 100L Teflon bag were: 1.5 - 2.8 ppm 2 - propoxyethanol, 1.07 - 2.07 ppm reference, 10 ppm CH_3ONO, and 2.5 ppm NO. These mixtures were allowed to sit for 30 - 60 minutes before background samples were taken. Samples were collected at 30 ml/min. on a Hastelloy trap at -86 °C for 4 minutes and

then flash injected at 260 °C onto the GC column. The range of photolysis times was 5 - 20 seconds for a combined time of no greater than 90 seconds. The maximum photolysis time corresponded to a decrease of 47% of 2-propoxyethanol. The FID signal was used to quantify the change in reactant concentrations.

An initial estimate for the 2-propoxyethanol / OH rate constant was made using the structural activity relationship (SAR).[3] Using this method, a rate constant of 24.8×10^{-12} cm^3 $molecule^{-1}$ s^{-1} was calculated. The most probable sites of attack by the OH radical were on either side of the ether oxygen ($CH_3CH_2\underline{CH_2}$ O $\underline{CH_2}CH_2OH$). Previous experimental work by Kerr gave a rate constant of 16.4×10^{-12} cm^3 $molecule^{-1}$ s^{-1} using 1-hexanol as the reference.[4]

Products from the OH reaction with 2 - propoxyethanol

The products from the relative rate experiments were confirmed using mass spectrometry. The ones that could be identified and were available commercially, were purchased.

Results

Rate coefficients of OH reaction with 2 - propoxyethanol

The runs conducted with nonane as a reference produced erratic data. Figure 1 shows the data. Differently shaped points indicate different runs. The correlation coefficient for this reference was poor. Since the rate constant for nonane was 10.3×10^{-12} cm^3 $molecule^{-1}$ s^{-1}, the rate constant for 2-propoxyethanol calculated from the slope was

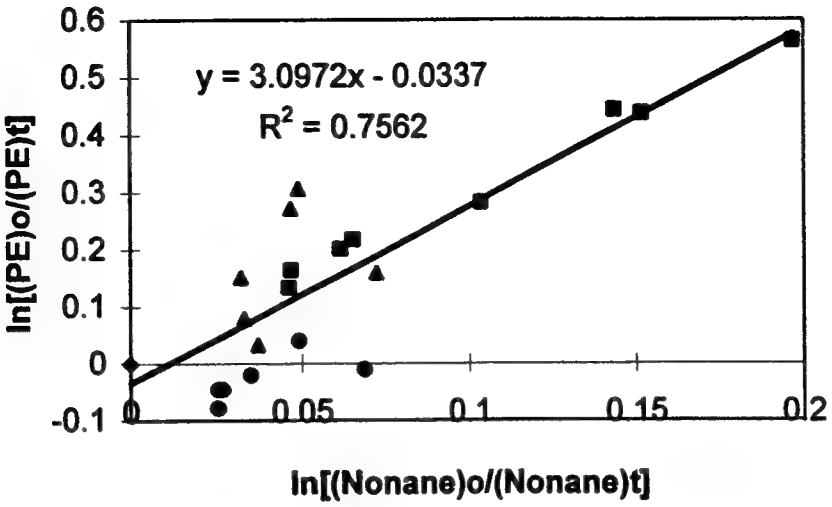


Figure 1: Rate Constant for PE/OH Reaction

31.6×10^{-12} cm^3 $molecule^{-1}$ s^{-1}. This was greater than the estimated rate constant by a factor of 1.27 and greater than previous experimental value by 1.93. The error involved in the data appeared to be systematic per day but random between days. As can be seen in Figure 2, linear least squares analyses of runs 2 and 3, performed on a Friday and the following Monday, respectively, produced different results. The individual correlation coefficients for each run were, however, quite good.

During the nonane runs, certain changes were made to the equipment to try and isolate the source of the error. The GC temperature program for the runs was changed from : $T_{initial}$ = 30°C , 3 min., ramp = 14°C/min., T_{final} = 150°C, 5 min. to: $T_{initial}$ = 35°C , 5 min., ramp = 10°C/min., T_{final} = 220°C, 5 min. This was done to increase peak separation. The program change had no effect. The mass spectrometer was then disconnected from the GC, leaving the FID as the sole detector.

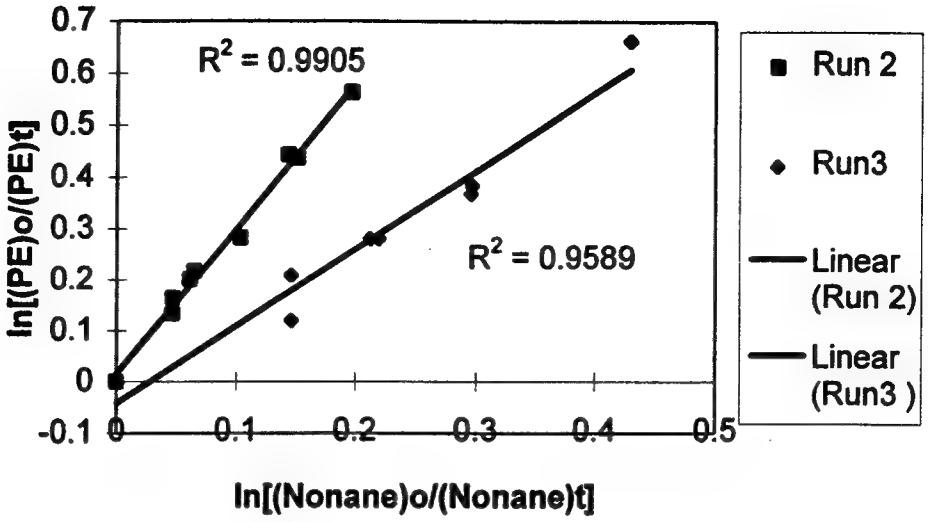


Figure 2: Rate Constant for PE/OH Reaction

This required changing the helium pressure in the FID from 25 psi to 2psi. This change was prompted by a concern that the split in the flow and the rather large guard column after the G.C. column may have been contributing to obstruction or uneven splitting of the flow. This change also had no effect. Finally, the line to the chamber and the trap were leak tested. No leaks were found.

The runs performed with hexane as the reference produced even more perplexing results. As can be seen from Figure 3, the data were nonlinear in nature. This suggested another loss mechanism for hexane that could not be attributed solely to OH radical reaction. Because of this inexplicable result, experiments proceeded using another reference.

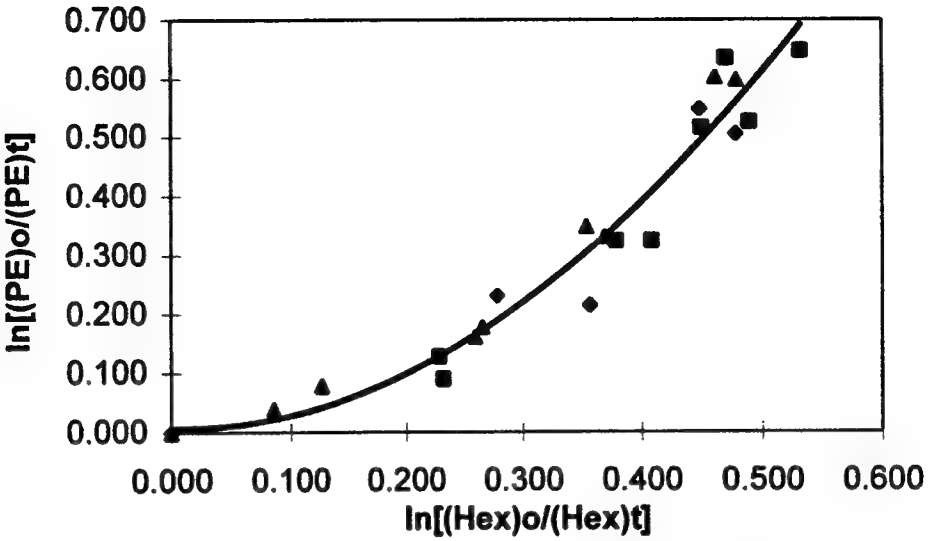


Figure 3: Rate Constant for PE/OH Reaction

The next set of runs was performed using dodecane as a reference. These data are plotted in Figure 4. Here, again there was a visible difference between runs performed on

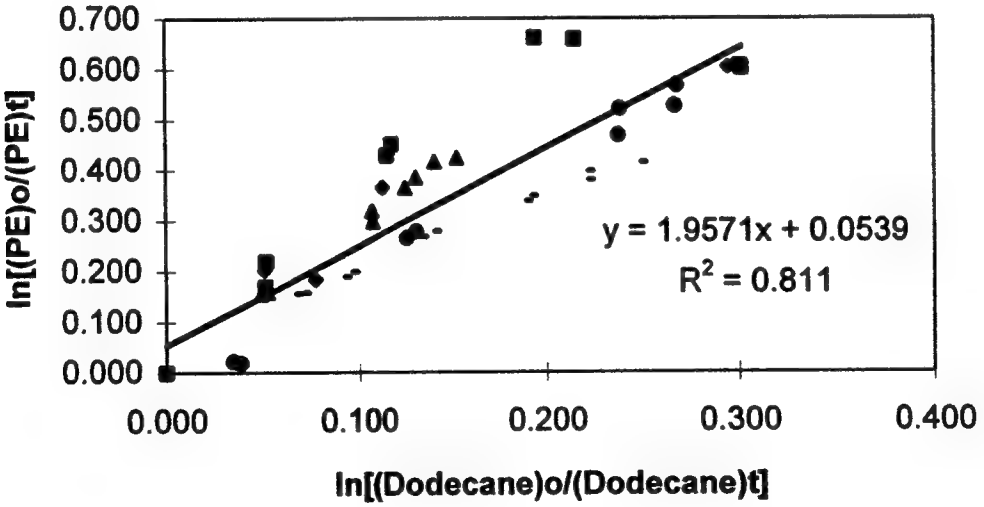


Figure 4: Rate Constant for PE/OH Reaction

different days. A few runs had slopes greater than the expected line and others had slopes below the expected line. Once again, individual runs had good correlation coefficients. The rate constant found for 2-propoxyethanol using the linear least squares analysis slope and dodecane's rate constant (13.9×10^{-12} $cm^3molecule^{-1}s^{-1}$) was $27.2 \pm 4.4 \times 10^{-12}$ $cm^3molecule^{-1}s^{-1}$. This value was greater than the estimated value by a factor of 1.10 and greater than the literature value by 1.66. The statistical error associated with the rate constant value was calculated using a 95 % confidence interval.

During the dodecane runs, more changes were made to the system. The rotary valve injector was completely replaced since the handle was hard to turn, and may have contributed to inconsistent injection. The new valve, the column, and the trap were all leak tested. No leaks were found. Finally, zero air from a cylinder replaced the "house air" supplied to the FID system. This was a concern due to the fact that when the house air was being used by other lab personnel, the air pressure in the FID went down, causing the FID signal to decrease. All of the changes that were made had no effect on the results.

The final set of runs was performed with 1-heptanol as the reference. As can be seen from Figure 5, the correlation coefficient for these data was quite poor. In the plot, differing slopes were more apparent than in the dodecane runs. The linear least squares analysis produced a slope of 1.39. This corresponded to a 2-propoxyethanol rate constant of $19.0 \pm 3.8 \times 10^{-12}$ $cm^3molecule^{-1}s^{-1}$, using the rate constant of

13.7×10^{-12} $cm^3 molecule^{-1} s^{-1}$ for 1-heptanol. This value was lower than the estimated rate constant by a factor of 1.3 but was greater than the literature value by a factor of 1.16. The error was calculated using a 95% confidence interval.

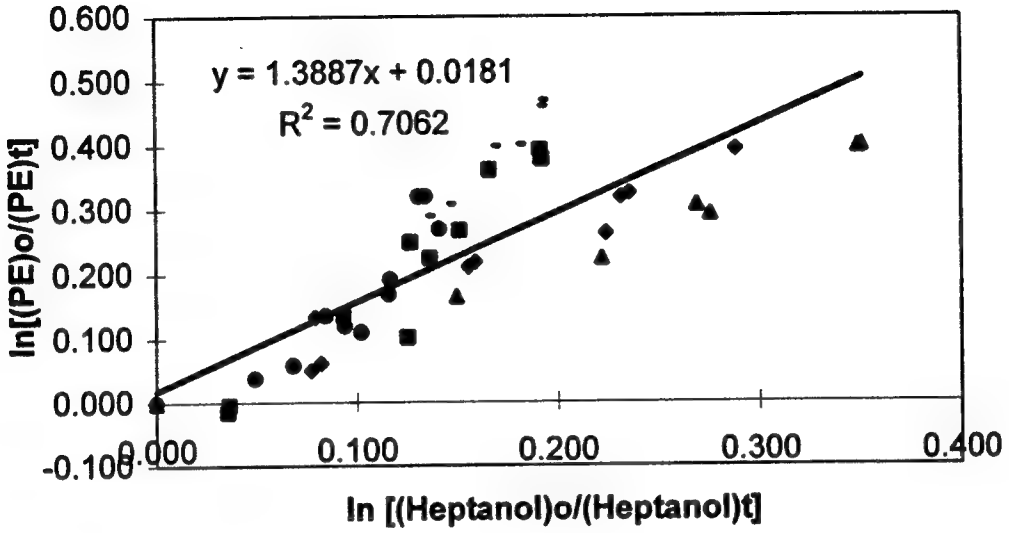


Figure 5: Rate Constant PE/OH Reaction

Products from the OH radical reaction with 2-propoxyethanol

The following products were found using mass spectrometry during the nonane runs: propyl formate ($HC(O)OCH_2CH_2CH_3$), acetaldehyde ($CH_3C(O)H$), propyl nitrate ($CH_3CH_2CH_2ONO_2$), 2-ethyl-1,3-dioxolane ($CH_3CH_2CHOCH_2CH_2O$), and 1,2-ethanediol monoformate ($HC(O)OCH_2CH_2OH$). The most observable product was propyl formate. This was expected as ethyl formate and butyl formate were found from previous studies done on 2-ethoxyethanol and 2-butoxyethanol, respectively. [5,6]

Conclusions

The data from the relative rate experiments suggested a systematic error was present which appeared randomly between days. Since many runs had apparently differing slopes, it was likely that compounds were sticking to either the bag or the trap and coming off at different times during the runs. This would have the effect of changing the x and y components of the slope. For example, if 2-propoxyethanol was sticking to the trap, and coming off at later times, this would increase the concentrations at time t, decrease the y component, and thus lower the slope. If the reference was sticking to the trap, its concentrations a time t would be higher, the x component would be lower, and thus the slope would rise.

The two final references, dodecane and 1-heptanol, were used to compare experimental rate constants. Taking into account both references' statistical error, the two rate constants do have a slight overlap. At dodecane's lowest value and 1-heptanol's highest value, the calculated rate constants do overlap at 22.8×10^{-12} $cm^3 molecule^{-1} s^{-1}$.

Due to the unreliable rate constant data collected, a full product study was not performed. Some products were identified by mass spectrometry. These included propyl formate, acetaldehyde, propyl nitrate, 2-ethyl-1,3-dioxolane, and 1,2-ethanediol monoformate. It must be noted that these products were identified during a relative rate experiment, where another reactant, the reference, was present. This would make it unclear as to which reactant the product came from.

Recommendations

The Hastelloy trap should be replaced. If active sites, where compounds become attached and possibly react, are present on the trap this will affect the measured concentrations. The compounds may be sticking to the Teflon bag so the bag might need to be changed as well. The next inevitable step would be to change the column. Finally, some rate constant studies with 2-propoxyethanol should be done on another system, to confirm that the problem is not due to the equipment used in this experiment.

The runs with dodecane and 1-heptanol should be redone once the system appears to be receiving consistent data. Also, a complete product study should be done once the rate constant data look reasonable.

Acknowledgments

My sincere appreciation to Air Team, Research & Development Laboratories, Air Force Office of Scientific Research, and Armstrong Laboratory, Tyndall AFB for the use of facilities and equipment.

Special Thanks to Dr. J.R. Wells, Dr. Jean Andino, Mr. Stewart Markgraf and Mr. Steve Baxely.

References

1. D.F. Smith, T.E. Kleindienst, E.E. Hudgens, and J.J. Bufalini, Intern J. Environ. Anal. Chem., 54, 265 (1994).

2. W.D. Taylor, D. Allston, M. J. Moscato, G. D. Fazckas, R. Kozlowski, and G.A. Takacs, Int. J. Chem. Kinet., 12, 231 (1980).

3. E.S.C. Kwok, and R. Atkinson, Atmos. Env., 29, 1685 (1995).

4. K. Stemmler, D. J. Kinnison, and J. A. Kerr, J. Phys. Chem., 100, 2114 (1996).

5. K. Stemmler, W. Mengon, J. A. Kerr, Environ. Sci. Technol., 30, 3385 (1996).

6. K. Stemmler, W. Mengon, D.J. Kinnison, J. A. Kerr, Environ. Sci. Technol., 31, 1496 (1997).

THE EFFECTS OF OBSERVATION AND TRAINING SCHEDULE ON THE ACQUISITION OF A COMPLEX COMPUTER BASED TASK

Julie A. Stiles-Shipley
Graduate Student
Department of Psychology

Bowling Green State University
Bowling Green, OH 43403-0001

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

THE EFFECTS OF OBSERVATION AND TRAINING SCHEDULE ON THE ACQUISITION OF A COMPLEX COMPUTER BASED TASK

Julie A. Stiles-Shipley
Graduate Student
Department of Psychology
Bowling Green State University

Abstract

The effects of observation and training schedule on the acquisition of complex computer based task are being studied using a 2 (Observation) x 2 (Training) x 2 (Experience) between subjects design. Participants consist of 240 males and females who are divided into 8 different experimental conditions. Participants either observe a model play several sessions of Space Fortress or they play Space Fortress without any observation. The skill level of the model varies as does the number of observation and practice trials. Average scores on the test sessions for Space Fortress will be analyzed in using a 2 (observation) x 2 (Training) x 2 (Experience) x 6 (Tests) mixed factors analysis of variance (ANOVA). It is expected that there will be a main effect of observation and a main effect of experience. In addition there may be an interaction between experience and training schedule and with observation and training schedule. Furthermore, a three way interaction between observation, experience and training schedule would not be unexpected. Therefore, we predict that observation can be useful for facilitating the training of a complex computer based skill once the proper training schedules and model characteristics are defined.

Introduction

Observation is important for learning in that it enables an individual to model behavior when learning a new task as well as shape behaviors to improve performance on an existing task. The social cognitive theory of observational learning (Bandura, 1986) supports that "learners extract information from modeled behaviors to form a cognitive representation of the actions that guide response production and serves as a standard for error correction through feedback" (Goettl & Connolly-Gomez, 1996, p. 1). Recently a series of studies (Connolly-Gomez & Goettl, 1995; Goettl & Connolly-Gomez, 1995; Goettl & Connolly-Gomez, 1996) have investigated the role of observational learning in Computer-Based Training (CBT). Computer-Based Training traditionally has been used to train individuals. However, by using observation as part of the training paradigm, the number of people who can simultaneously receive training can increase without increasing cost demands. Before a paradigm can be developed, several issues must be resolved to ensure the most efficient schedule possible.

Bandura's (1986) theory of observational learning offers a set of guidelines for when and how observation will best benefit performance for CBT. He contends that observational learning is determined by four things: attention, retention, behavioral production, and motivation. Specifically, participants must be able to attend and to extract important information from the task through visual processes. Visual interpretations are transformed into some form of cognitive representation of the task. The cognitive representations serve as a model of the task which participants can refer to at any time. According to Carroll and Bandura (1982), encoding and rehearsing symbolic codes is critical to the formation of cognitive representations for guiding performance. Without the proper representations, performance will be flawed due to improper transformation of observed action into symbolic codes. During the behavioral production stage, participants translate their cognitive representations into actions necessary for performing the task. Actions can be changed by comparing the actual results to the cognitive representations and making the necessary modifications to achieve desired performance. This continues until a close correspondence between the internal representation and the action is achieved (Carroll & Bandura, 1985). Finally, motivation is used to govern the behavioral production process. "Motivational processes influence observational

learning by affecting selective attention and symbolic coding and rehearsal. Insufficient incentives will also have a more direct effect on performance itself" (Carroll & Bandura, 1982, p. 165) Overall, if the features of a task can be easily extracted to form cognitive representations, then observation can serve as a means of skill acquisition.

Several studies have shown that observation can be a useful tool for training cognitive (Berry & Broadbent, 1984; Berry, 1991; Goettl & Connolly-Gomez, 1996; Shebilske, Regian, Arthur, Jordan, 1992) as well as motor tasks (Blandin, Proteau, Alain, 1994; Carroll & Bandura, 1982; Carroll & Bandura, 1985; Lee & White, 1990; McCullagh & Caird, 1990). Shebilske et al. (1992) examined an Active Interlocked Modeling (AIM) protocol which paired trainees so that they could combine observation with hands-on practice on a complex task. Each person performed a separate component of a complex task (Space Fortress) simultaneously. As a result, participants received feedback from their partners performance as well as their own. They could then apply and practice this knowledge when their roles were reversed. Performance did not decline as a result of decreased practice time, indicating that subjects learned from observing their partners. In a follow up study, Shebilske, Jordan, Arthur, and Regian (1993) increased the dyads to four. They found that four trainees could learn a task as well as one which decreased hands-on performance, training time, and resources by seventy five percent. However, not all tasks may be suitable for such a training protocol.

Vision has been shown to be extremely important in the learning and regulation of complex motor tasks (Posner, Nissen, & Klein, 1976; West, 1967). If important characteristics of a task cannot be visually monitored then the observer will be at a serious disadvantage when trying to imitate the action (Carroll & Bandura, 1982). Therefore, observation alone is restricted to the interrelationships of task variables, such as the cognitive and strategic components of the task, but it is not useful for learning how to control the system itself (Funke & Muller, 1988, as cited in Berry, 1991). In an experiment that examined the role of action in implicit learning, Berry (1991) found that subjects could not learn to control a task simply by observing an individual's interaction with the task. Berry explained this by classifying the task as one which involved implicit learning. Specifically, the actions and decisions used to control the task were

not visually obvious and could not be verbalized. The relationship between actions and decisions in this experiment were relatively non-obvious, thereby making it difficult for an observer to form a useful symbolic representation of the task. Berry classifies this type of task as non-salient. A non-salient task is one in which "the relationship between input and output variables is not obvious, and there is a dissociation between ability to perform and verbalizable knowledge" (Berry, 1991, p.890) . Tasks are considered salient if "the relationship between input and output variables is relatively obvious, and for such tasks the ability to perform agrees with verbalizable knowledge" (1991, p. 890). It is possible then that the salience of the task will determine whether the task can be learned through observation. Currently, there is not an independent way of defining the salience of a task (Berry, 1991).

Goettl and Connolly-Gomez (1995) introduced CBT tasks that included cognitive as well as psychomotor functions to explore the conditions under which observational learning is an effective pedagogy for computer-based flight simulator tasks. They hypothesized that observational learning would have more of a benefit for tasks that could be symbolically coded than for tasks that cannot. Specifically, performance on tasks that have high cognitive and strategic demands will be better than performance on tasks that require more perceptual-motor responses. To test this idea, they used two versions of a flight simulator task; one which required participants to fly through gates positioned at various locations in space (Slalom Task) tapping perceptual-motor responses, and one in which participants navigated through open space to destroy three targets (Strike Task) tapping cognitive-strategic demands. As expected they found that observation had no effect on the Slalom Task but did have some benefit for the Strike Task. Performance on the Strike Task was superior for those participants who received observation than for the controls. However, only female participants showed this effect which raises additional questions regarding observational learning.

Goettl and Connolly-Gomez (1995) noted the differences in skill level of the males and females on the Strike Task as one explanation for the discrepancy. Because males achieved a higher skill level than females, it was more likely that female observers watched someone with a higher skill level perform the task. Therefore, observation may only

benefit learning of a task if the person being observed is of a higher skill level than the viewer. This notion is consistent with Anderson's (1983) theory of skill acquisition which stresses that exposure to the target task is necessary in order to develop skills in their proper context. In other words, before an individual can form the proper cognitive representation of a task he/she must first view the task under the desired conditions. Bandura suggests that "the structure of complex behavior can be symbolically constructed more readily from observing the behavior displayed in an already integrated form than in trying to construct it gradually from examining the effects of one's exploratory actions (Carroll & Bandura, 1985, p. 280). An alternate explanation involves the stages of learning during which observation is most beneficial.

During the early stages of learning, information is believed to be in the form of declarative knowledge. Declarative knowledge relies heavily on explicit information and feedback. Because symbolic representations are often declarative and observational learning relies heavily on symbolic representations, it makes sense that observations should be made during this stage of learning. As a person becomes more skilled, information becomes more automatic and he/she relies less on explicit knowledge. It may be that observation is most effective during the declarative stages of knowledge, when behavioral productions are more controlled. Berry (1991) investigated whether performance would improve in the absence of action over an extended period of time. She found that observations must be tied to action in order to be effective especially in the early stages of learning. Therefore, if males become skilled faster than females, they will progress through the early stages of learning more quickly and benefit less from observation. As a result, observation may only benefit performance if it is presented in the early stages of skill acquisition along with practice opportunities. However, because Goettl and Connolly-Gomez (1995) did not examine this notion in their experiment, further investigation into this topic is necessary before any conclusions can be drawn.

Blandin et al. (1994) reported that one factor thought to influence the cognitive processing of an individual learning a new task is their practice schedule. It has been shown that a motor skill is best learned by using a blocked (massed) practice schedule rather than a random practice schedule (spaced), but the opposite is true for

performance on retention tests. This is known as the contextual interference effect. There are two explanations for contextual interference (CI), elaboration distinctiveness and action plan reconstruction. Elaboration distinctiveness is based on the idea that different movement patterns to be learned are stored in short term memory. Randomizing the presentation of trials forces an individual to identify which of the stored movement patterns is appropriate to complete the trial. Therefore, "there is more practice in a random condition which allows for more opportunities to compare movement patterns and the more distinct they become from one another" (Blandin et al., 1994, p. 18). Action plan reconstruction on the other hand, suggests that "parts of a particular movement pattern may be forgotten between trials and that the subject must actively reconstruct the required action plan before initiating a particular trial" (p. 19). Even though the CI effect has been established in a variety of motor tasks (see Magill & Hall, 1990, for review), it does not always exist in cases which involve observational learning.

Lee and White (1990) tested the hypothesis that if an observer engages in the same cognitive activities as the models, then observers should exhibit a similar CI effect as the models for blocked and random schedules of practice. In other words, observers who view models in a random practice schedule should perform better on retention tests than observers who view models in a blocked practice schedule. Lee and White (1990) found that observers performed equally well on retention tests regardless of practice schedule and that the beneficial effects of observation can overcome the disadvantages usually found for a blocked practice schedule. In a replication experiment, Blandin et al. (1994) also showed that there were no differences between the blocked and random practice schedules for participants who observed a model's performance. Goettl and Connolly-Gomez (1996) addressed some of these issues by examining the differences between massed and spaced observation schedules.

Goettl and Connolly-Gomez (1996) hypothesized, based on Anderson's (1983) theory of skill acquisition, that massed observation would be better than spaced observation because it allows for a richer development of a mental representation of the task prior to hands-on practice. As predicted the massed observation groups performed better than the spaced-observation groups. Furthermore, participants in the

spaced observation condition did not benefit from observation possibly because massed-observation participants observed players with higher skills prior to their first training session. Moreover, in the massed-observation performance condition, participants had the opportunity to watch performers overcome their mistakes and develop effective strategies for the task, whereas in the spaced-observation condition actual performance always preceded observation. There are two explanations as to why the massed condition outperformed the spaced condition.

The first explanation suggests that it is necessary for a person to observe an individual progress through the stages of learning and to watch as the individual becomes more skilled at the task. This allows the observer to experience common mistakes that occur early in skill acquisition and to avoid those mistakes during his/her own performance. By observing a model for an extended length of time before having to perform, the observer can witness the model's progression from novice to experienced without interruption. Therefore, it is possible that a higher skill level can be reached by those who practice continually and as a result, observers have the opportunity to view some skilled performances of the model. If a higher skill level is reached through massed practice then observers have an advantage of seeing the task performed in its proper context. The second argument that explains why massed observation is better than spaced observation relies on this assumption.

As previously suggested by Carroll and Bandura (1985) individuals will learn more about a task if they view it in a form that is already integrated. Skilled models are ones who have learned to properly integrate the cognitive and motor components of the task in order to achieve optimal performance. Therefore, observers can witness the task in an integrated form and in the proper context before they are required to perform. Observing a task in the proper context allows individuals to set goals for their own performance and teaches them how to organize information for the task. Regardless of which explanation is correct, it would seem that when trying to train a person on a task using observational learning, the skill level of the model is an important factor to consider. However, because the Goettl and Connolly-Gomez (1996) subjects benefited from massed practice as opposed to spaced practice using novice models. It may be the case that the best type of

practice schedule to use for training will depend on the skill level of the model. Therefore, it is necessary to examine the relationship between practice schedule and skill level. In the proposed experiment, two variables that will be manipulated are skill level of the model and practice schedule.

The proposed experiment will examine the effectiveness of observational learning for a complex computer-based skill as a function of observed skill level and training schedule. The three main research questions to be addressed are: 1) will people benefit more from observation if the model they are observing is experienced as opposed to a novice, 2) will a massed or spaced training schedule facilitate observational learning, and 3) are different training schedules necessary for observing different skill levels? Previously, Goettl and Connolly-Gomez (1996) found that males did not benefit from observational learning in the Phoenix task whereas females did show an effect. One explanation suggested that males were more skilled at performing the task than females and therefore females were able to observe performers of a higher skill level more often than males. It seems that individuals may only benefit from observation if the model whom they are observing is of a higher skill level. Higher skilled models offer the opportunity for observers to see the task being performed correctly or in the proper context which allows observers to form a target or a goal for their own performance. Anderson's (1983) theory of skill acquisition is consistent with this notion. Our first hypothesis will examine the effects of observed skill level on performance. We predict that observation can facilitate the acquisition of a new skill as long as the model performing the task is more experienced at the task than the performer.

The practice schedule used for observational learning also affects the skill level of the model and the observer. Typically in training, and consistent with theories of contextual interference, spaced practice schedules produce better learning than massed practice schedules. However, observational learning may be one case in which massed practice is more beneficial to skill acquisition than spaced practice. In the massed practice condition, participants can observe performance for longer periods of time so that they have time to observe the strategies being used and to form the proper cognitive representations of those strategies. The ability to form a representation of a task is critical

for observational learning to occur. In spaced practice, short intervals of observation may not be enough for an individual to extract the proper strategies from the model's performance. Therefore, it may take several sessions before any representations of the task are formed. In addition, massed practice allows the models to become more highly skilled because of the continuous practice, and it allows the observers more time to practice several strategies at one time. The second hypothesis to be tested will examine the effects of training schedule on performance. We predict that participants in the massed training schedules will perform better than participants using spaced training.

If models are already skilled at the task, then they do not need the extra time in massed practice to improve performance. Therefore, spaced training may be sufficient for observational learning if the model is highly skilled. Our third hypothesis will test if there is an interaction between observed skill level and training schedule. We predict that if massed training allows a performer to become more highly skilled, then massed training will benefit participants observing a novice model. Moreover, participants who observe a skilled model will benefit more from spaced training because they have the opportunity to practice the proper strategies shortly after observing them. In pursuit of these objectives, we propose the following study.

Methodology

Participants

Two hundred fifty two males and females will serve as participants in this experiment. Fifteen males and 15 females will participate in the massed and spaced training schedule conditions with either a novice or experienced model. In addition 15 males and 15 females will make up the control conditions for the two training schedules. The experienced and control models will compose the two all-perform control conditions that consist of 30 participants each. All participants will be recruited by local temporary employment agencies and were paid approximately $5.00 per hour for their participation. Participants ages will range from approximately 18 to 30 years. All participants will have at least a high school diploma or Graduate Equivalency Degree (GED). Participants will be screened to eliminate those who had previous experience with Space Fortress or who reported spending more than 20 hours a week playing video games, with the exception of the experienced models. The

experienced models will be chosen according to previous performance scores on Space Fortress.

Materials

The data will be collected in a laboratory containing 23 Compaq DeskPro 486/33L computers with NEC/Multisync SVGA monitors and CH Products Flight Sticks.

Tasks

Space Fortress. We will use the version 4.5 of the Space Fortress task that is most similar to the version used by Gopher, Weil and Bareket (1994). In Space Fortress task subjects will use a joystick to operate a spaceship that is moving in a frictionless and hostile environment. The spaceship is being threatened by the Fortress positioned in the center of the screen and by mines flying around the screen. The goal of the task is to circumnavigate the Fortress while avoiding missile fire, to defend the ship from the mines and ultimately, to destroy the Fortress. There are several additional factors that must be identified during the game in order to achieve optimal performance. Subjects will need to identify whether a mine is a friend or a foe because a different course of action will be applied depending on the type of mine. Additionally, the subjects need to be prepared for the selection of bonus opportunities when they appear on the screen. The four sub-scores will be presented on the screen, which include Speed, Velocity, Points and Control; the summation of these sub-scores constitutes the total score in the task.

Cognitive Abilities Tests (CAM). We will use the version 4.0 of Cognitive Abilities Tests. The subjects' spatial processing speed, working memory, declarative fact learning and procedural skill learning will be assessed through the CAM-4 battery. The tests will be completed using the same computers as the ones to be used for the Space Fortress task. Cognitive Abilities Tests are described in more detail by Kyllonen, Woltz, Christal, Tirre, Shute, and Chaiken (1990).

The Uncued Simple Reaction Time (USRT). We will use the USRT task to record the arousal level in subjects. The participants will be required to press the down-arrow key on the keyboard whenever a red square appears on the screen and reaction time will be recorded in milliseconds. The occurrence of red square will be random and unexpected. We will administer the USRT task at the beginning of each

day, before lunch, and before the final session on each day to assess arousal level throughout training.

The Confidence and Alertness Questionnaire (CAQ). We will use the CAQ (Corrington, 1994) to rate different aspects of subjects' confidence and alertness with regards to their performance on the Space Fortress task. The CAQ contains 12 questions which ask the participants about their confidence and knowledge of the rules and strategies, their performance in comparison to other players, their ability to improve their performance, alertness and motivation to perform, and their thoughts about the task and its effect on performance.

Strategy Questionnaire. We will use a strategy questionnaire that was modified from the Learning Strategy Inventory (Dansereau, Long, McDonald, & Atckinson, 1975) to identify a set of strategies that subjects may find to be beneficial to their performance on the Space Fortress task. Responses to questions will be measured using a 6-point Likert scale. The questions will ask subjects to what extent a particular strategy was helpful or if they used a particular strategy while performing the Space Fortress task.

Windows Solitaire and Mine Sweeper. Microsoft Solitaire and Mine Sweeper games will be used as a filler task for the control groups in both, massed and spaced conditions.

Procedure

Testing will take place over two and a half days. All groups will follow the same general procedure for testing although modifications will be made depending on the specific condition. The general procedure is as follows: On day one testing will take place in the afternoon only, and all participants will complete an aiming task that will be used as a qualifying task. Anyone who does not reach the minimum score of 780 points will be dismissed from the experiment. Participants will then receive written instructions on how to play Space Fortress and watch a short demonstration video. Upon completion of the instructions, all participants will play four games of Space Fortress that will serve as their pretest measure. Following the pretest, participants will complete the CAM battery. The CAM battery consists of two tests to assess cognitive and spatial abilities. Once the CAM tests are complete, participants will be dismissed for the day. Training on Space Fortress will occur on the second and third days. Only one type of training

schedule will be used per week and each week will alternate between massed and spaced training. Each condition's protocol will be discussed separately.

Massed Training Condition

Models. Before playing Space Fortress, models complete a Cognitive Arousal Questionnaire (CAQ) and an Uncued Stimulus Reaction Time (USRT) test to measure the models alertness. Participants will then receive a written summary of instructions along with a description of the optimal strategies for playing Space Fortress while watching a summary video containing the same information. Models will play three sessions of Space Fortress. Each training session consists of 10 games and each game lasts for three minutes. After each session, participants will receive a 10-minute break. Upon completing the third session, models will complete a Performance Strategy Questionnaire (PSQ) which will investigate the strategies the models used while playing. Following the PSQ, models will perform a test session consisting of four Space Fortress games to asses the models' performance after three training sessions. Once again the model will then complete the CAQ and the USRT. The same procedure will be repeated for another three training sessions.

The third day will follow the same procedure as above except for the last test session. The last testing session is composed of several retention and transfer tasks. The first two games played will serve as a retention test. The first retention test is followed by a transfer task in which participants must tap their feet on a tapping board beneath their stations, alternating left and right. Two games will be played under these conditions. The next retention test will include two games of Space Fortress followed by another transfer test. In the second transfer test, participants must play two games of Space Fortress using the keyboard instead of the joystick. The final test consists of two games of Space Fortress which measures retention. Once models have finished all of the retention and transfer games, they will complete the PSQ, the CAQ, and an implicit/explicit (I/E) questionnaire. Models will be dismissed at the conclusion of the I/E questionnaire.

Observers. Participants who are in the observation condition will follow the same procedure as the models except for the first three training sessions on days two and three. Instead of playing Space Fortress during the first three sessions, observers will watch either an experienced or novice model play Space Fortress. At the end of the third

session, observers will complete a model strategy questionnaire (MSQ). The MSQ contains questions similar to the PSQ, however, the questions refer to the model's strategies and performance for Space Fortress. For the next three training sessions, participants will then play Space Fortress and follow the same procedure as the models. The same procedure is repeated on day three as day two. The observers will receive the PSQ at the end of training on day three so that they can identify their own strategies for playing Space Fortress. All retention and transfer tasks will be the same as the ones used for the models.

Controls. Participants who are in the control condition will follow the same schedule as the observer except for the observation. Alternating games of Solitaire and Mine Sweeper will be played in replacement of observation. Therefore, while the observers watch their models play Space Fortress, the control subjects will play either Solitaire or Mine Sweeper. At the end of three sessions, control subjects do not complete a strategy questionnaire but rather continue to play Solitaire or Mine Sweeper for the length of time it takes the models and observers to finish their questionnaires. All other retention and transfer tasks will follow the same procedure as the models and the observers.

Spaced Condition

Models. Before playing Space Fortress, models complete a Cognitive Arousal Questionnaire (CAQ) and an Uncued Stimulus Reaction Time (USRT) test to measure the models alertness. Participants will then receive written summary instructions along with a description of the optimal strategies for playing the Space Fortress while watching a summary video containing the same information. Participants will keep the written instructions at their stations for further reference. Models will play a total of six sessions of Space Fortress per day. Each training session consists of 10 games and each game last for three minutes. After each session, participants will receive a 10-minute break. Upon completing the sixth training session, models will complete a Performance Strategy Questionnaire (PSQ) which will investigate the strategies the models used while playing. Following the PSQ, models will perform a test session consisting of four Space Fortress games to asses the models' performance after three training sessions. Once again the model will then complete the CAQ and the USRT. The same procedure will be repeated for another three training sessions.

The third day will follow the same procedure as above except for the last test session. The last testing session is composed of several retention and transfer tasks. The first two games played will serve as a retention test. The first retention test is followed by a transfer task in which participants must tap their feet on a tapping board beneath their stations, alternating between the left and right foot. Two games will be played under these conditions. The next retention test will include two games of Space Fortress followed by another transfer test. In the second transfer test, participants must play two games of Space Fortress using the keyboard instead of the joystick. The final test consists of two games of Space Fortress which measures retention. Once models have finished all of the retention and transfer games, they will complete the PSQ, the CAQ, and an implicit/explicit (I/E) questionnaire. Models will be dismissed at the conclusion of the I/E questionnaire.

Observers. The procedure for the observers is similar to the models' procedure in that all testing sessions and questionnaires will be administered at the same as previously mentioned. However, participants who will be in the observation condition will observe and play alternating games of Space Fortress within a given session. Alternating games will be played in sets of two, meaning that participants in the observation condition will watch a model play two games of Space Fortress and then return to their own computer station to play two games of Space Fortress. This is repeated until a total of 10 games have been completed. In the first session, observers begin by watching a model play two games of Space Fortress. Therefore at the end of session one, participants have observed a total of six games and played a total of four games. Ten minute breaks are given in between each session. The same procedure is followed for the next six sessions. However, the total number of games played alternate from either six to four . At the end of the sixth session, observers will complete a model strategy questionnaire (MSQ). The same procedure is repeated on day three as day two. The observers will receive the PSQ at the end of training on day three so that they can identify their own strategies for playing Space Fortress. All retention and transfer tasks will be the same as the ones used for the models.

Controls. Participants who are in the control condition will follow the same schedule as the observers except for the observation. Alternating games of Solitaire and Mine Sweeper will be played in

replacement of observation. Therefore, while the observers watch their models play Space Fortress, the control subjects will play six minutes of either Solitaire or Mine Sweeper. However, at the end of the sixth session, control subjects will not complete a strategy questionnaire but rather they will continue to play Solitaire or Mine Sweeper for the length of time it takes the models and observers to finish their questionnaires. All other retention and transfer tasks will follow the same procedure as the models and the observers.

Predictions

Observation should facilitate the acquisition of Space Fortress because participants will have an opportunity to observe the game before they are expected to play. Furthermore, it is predicted that the experience level of the model will affect the quality of the information the observers see which in turn will influence their learning of the game. Those who observe experienced models will perform better using a spaced training schedule, whereas, observing a novice model will have more of an effect if they use a massed training schedule. Alternatively, observation may only be helpful in massed schedules and be ineffective in spaced training schedules despite the skill level of the model.

References

Anderson, J.A. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Bandura, A. (1986). *Social foundation of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall

Berry, D.C. (1991). The role of action in implicit learning. *The Quarterly Journal of Experimental Psychology, 43,* 881-906.

Berry, D.C. & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology, 36A,* 209-231.

Blandin, Y., Proteau, L., & Alain, C. (1994). ON the cognitive processes underlying contextual interference and observational learning. *Journal of Motor Behavior, 26,* 18-26.

Carroll, W.R. & Bandura, A. (1982). The role of visual monitoring in observational learning of action patterns: making the unobservable observable. *Journal of Motor Behavior, 14,* 153-167.

Carroll, W.R. & Bandura, A. (1985). Role of timing of visual monitoring and motor rehearsal in observational learning of action patterns. *Journal of Motor Behavior, 17,* 269-281.

Connolly-Gomez, C., & Goettl, B.P. (1995). Investigating tradeoffs between practice and observation in automated instruction. *Proceedings of the Human Factors Society & Ergonomics Society 39th Annual Meeting, 2,* 1340-1344.

Corrington, K.A. (1994). Spacing effects in the acquisition and maintenance of a complex skill. Unpublished master's thesis, TX A&M University, College Station, TX.

Dansereau, D.F., Long, G.L., McDonald, B.A., Actkinson, T.R. (1975). Learning strategy inventory development and assessment. Technical Report AFHRL-TR-75-40, Brooks Air Force Base, TX.

Funke, J., & Muller, H. (1988). Eingreifen und prognostizieren als Determinanten von System-identifikation und Systemsteuderung. *Sprache und Kognition, 7,* 176-186.

Goettl, B.P., & Connolly-Gomez, C. (1995). The role of observational learning in automated instruction of complex tasks. *Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society, 2,* 1335-1339.

Goettl, B.P., & Connolly-Gomez, C. (1996). The effects of observational learning and practice schedule on the acquisition of a complex skill. *Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: Human Factors and Ergonomics Society.

Goettl, B.P., Yardick, R.M., Gomez, C.C., Regian, J.W., & Shebilske, W.L. (1996). Alternating task modules in isochronal distributed training of complex tasks. *Human Factors, 38,* 330-346.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors, 36,* 387-405.

Kyllonen, P.C., Woltz, D.J., Christal, R. E., Tirre, W.C., Shute, V.J., & Chaiken, S. (1990). CAM-4: Computerized battery of cognitive ability tests. Unpublished computer program, Brooks Air Force Base, TX.

Lee, T.D., & White, M.A. (1990). Influence of an unskilled model's practice schedule on observational motor learning. *Human Movement Science, 9,* 349-367.

Magill, R.A., & Hall, K.G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science, 9,* 241-289.

McCullalgh, P., & Caird, J.K. (1990). Correct and learning models and the use of the model knowledge of results in the acquisition and retention of a motor skill. *Journal of Human Movement Studies, 18,* 107-116.

Posner, M.I., Nissen, M.J., & Klein, R.M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83,* 157-171.

Shebilske, W.L., Jordan, J.A., Arthur, W.Jr., & Regian, J.W. (1993). Combining multiple emphasis on components protocol with small group protocols for training complex tasks. *Proceedings of the 37th Annual Meeting of the Human Factors and Ergonomics Society (1296-1300).* Santa Monica, CA: The Human Factors and Ergonomics Society.

Shebilske, W.L., Regian, J.W., Arthur, W.Jr., & Jordan, J.A. (1992). A dyadic protocol for training complex skills. *Human Factors, 34,* 369-374.

West, L.J. (1967). Vision and kinesthesis in the acquisition of typewriting skills. *Journal of Applied Psychology, 51,* 161-166.

INTEGRATING MULTISENSORY DISPLAYS FOR AN ADAPTIVE TARGET LEADING INTERFACE

Robert S. Tannen
Department of Psychology

University of Cincinnati
ML 376
Cincinnati, OH 45220-0376

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

September 1997

INTEGRATING MULTISENSORY DISPLAYS FOR AN ADAPTIVE TARGET LEADING INTERFACE

Robert S. Tannen
Department of Psychology
University of Cincinnati

Abstract

An important research issue in the advancement of tactical cockpit interfaces is the determination of optimal mappings between information display and the perceptual capabilities of operators. Rapid and accurate acquisition of target positions is mission critical in these dynamic, multi-task environments, but the human factors which determine the success of situational information portrayal have not been thoroughly exploited. A promising approach to spatio-temporal data presentation is the use of multisensory interfaces that take advantage of natural perception and action behaviors. This study will collect target tracking, course deviation, and task workload measures for several different configurations of separate and integrated multi-sensory (head-mounted-visual and spatialized audio) target position displays during immersive threat localization/flight navigation simulations. Certain display configurations make use of an adaptive algorithm which will switch information delivery between visual and auditory channels depending on the target's current position with respect to the pilot's field-of-view (FOV). It is expected that the adaptive interfaces will improve target localization rate and efficiency over fixed single channel displays, comparable to the improved performance found with redundant multisensory displays. The multisensory displays should also reduce flight course deviations as pilots will be able to attend visually to boresighted navigational displays while information about targets beyond FOV will be displayed aurally. The anticipated performance benefits generated by situation-specific information portrayal are also expected to manage pilot workload. The results will lead to guidelines for the effective implementation and use of multisensory spatial information displays, as well as adaptive interfaces in general.

INTEGRATING MULTISENSORY DISPLAYS FOR AN ADAPTIVE TARGET LEADING INTERFACE

Robert S. Tannen

Introduction

The design of effective interfaces is dependent upon optimizing the relationship between information display and users' perceptual capabilities. This is an imperative issue within the information rich and space limited,environments of tactical aviation. Pilots must often take action based on the rapid integration of a number of complex system, situation, and mission-level factors. This research is focusing on the display of information for target localization and tracking. Target localization refers not only to finding threats, as this study emphasizes, but can be generalized to locating navigational landmarks, landing sights, spotting indicators of weather conditions, and a number of other out the window (OTW) events. Localization is a part of many everyday behaviors such as driving, but the mission-critical nature, large coverage areas, and compressed time-scales in tactical aviation necessitates the optimization of this kind of task performance.

Unaided target localization has been predominately performed via the high spatial acuity of the visual system. Aided target tracking has also relied upon optical information display. While people rely greatly on their visual system for the detection and examination of objects in space, it is often the auditory and tactual (at close range) systems which initially detect information about the presence and location of an object or event. Vision and audition are intimately tied; it is natural for a person to hear an event beyond the field of view, and then orient one's vision, based on that initial information, to extract additional, often more detailed, information. This tight perceptual link stems from the development of multisensory neurons that integrate information from the auditory and visual systems. Developmentally, the tonotopic visual system serves as the reference for the mapping of spatial auditory input (Stein, Meredith, & Wallace, 1994). Multisensory neurons promote perception-action cycles through motor connections which coordinate eye and body movements towards stimuli. The synergy of the perceptual systems, in light of the high visual task demand of the cockpit environment and the limited visual channel capacity of the pilot (Perrott, Cisneros, McKinley, & D'Angelo, 1996), suggest that it would be beneficial to develop virtual displays that take advantage of our natural, multisensory localizing capabilities.

Determining the characteristics of a multisensory interface goes beyond merely balancing information across the senses, for some forms of information may be more readily received through a preferred modality. Furthermore, certain information may be better delivered to one perceptual system at a given time, and, depending upon both environmental and operator conditions, other systems at other times (Cutting, 1986). For example, information about the current location of an object moving in space may best be described aurally during head movements, but when the head

is relatively static (between movements) information delivery may be best achieved through visual representation. The variability inherent in such dynamic situations suggests the incorporation of plasticity into interface schemes. Adaptive interfaces are capable of real time, operator and context dependent display changes (Bennett, 1997). A system that is capable of real-time changes in the modality of information display could have numerous benefits including improved operator performance, workload manageability, and enhanced situational awareness. This could be demonstrated by the simple paradigm relating perceptual system constraints to object spatial location, and could serve as the basis for an improved cockpit-based target locating system. A system could present information about targets to the auditory channel when beyond the visual field, and then to the visual channel when the spatial relationship between the operator and the target accommodates to the field of view.

The most concrete advantage such a display system is that a pilot would be to attend to relevant information both within and beyond his field of view. He could track the position of a threat aurally while looking at a weapons status display in order to determine how to deal with this threat. Compare this with having to follow the changing positions of a target on one screen, and switch fixations to focus on another screen, to check weapons status, and the potential for enhanced performance and reduced workload should become apparent. This type of interface would be especially strong for target detection because a pilot could potentially track multiple targets in distal fields of view by using his visual and auditory senses.

The success of such an adaptive, multisensory interface a system rests on exploiting the human-environment relations (affordances) that vary across modalities, and this will require addressing several significant research questions. The relationship between the pilot's field of view (FOV) and the target provides a measurable, invariant, and direct limitation on the utility of the visual system. This constraint can be used to drive an adaptive algorithm that could switch information display over the modalities. Investigation will be required to determine what kinds of changes should be driven by this algorithm. Modulating between single modality displays might seem advantageous because a pilot would receive information through the most appropriate channel, freeing up the other sense to work on another task. Alternatively, target tracking performance may decline, depending on additional tasks, because of divided attention (Moroney, 1997). Single modality display also lacks the signal strength of simultaneous multisensory stimulation (Stein, et al, 1994).

As the physical/perceptual restrictions of the field of view limit the functional range of the visual system, it is worthwhile to consider other constraints on the perceptual systems which could guide display usage. Note, that although virtual displays are not limited to normal constraints, they must ultimately provide information in a form that is perceivable. For example, a simulated light ray could turn around corners, or a sound could appear to travel faster than the speed of sound, but the perceptions of such events are ambiguous. The significant differences between the acoustical

and optical conveyances of information will now be discussed in detail, starting with global features and moving to specific considerations for optical/acoustical interface design.

The optical and acoustical components of an event, such as a passing target, can provide different information in both form and content. Some literature has ascribed the properties of optics, to characterize acoustics. This includes Lee's (1990) description of an acoustic array, similar to an optical array, that provides a multitude of potential listening points. Each point is defined by the product of the waveforms that, through direct and indirect paths, converge at the location at a given time. Information about the environment is singular and potentially attainable at each point. Although Lee's ideas about point specificity probably hold, his analogy with optics oversimplifies the underlying acoustics that generate individual sound spaces. Gibson (1966) declared, "ambient sound is not even comparable to ambient light". Therefore, it is useful to distinguish between the properties of optics and acoustics to minimize the need to rely on ecological optics in the determination of how acoustical characteristics may be used in interface design.

The distinctive characteristics of electromagnetic and mechanical energies determine the type and form of information that can be carried through light and sound, respectively. An important difference is in the temporal constraints. The speed of light is orders of magnitude greater than the speed of sound. Information about an event, and changes over time, can be much more rapidly conveyed through optics than acoustics. For proximal distances this speed advantage is trivialized due to the relatively slow processing speed of the human perceptual system. For more distant events the speed difference can be obvious, for example a flash of lightening preceding the clap of thunder. Notably, the temporal relationship between the arrival of the optical and acoustical components of the same event is a potential invariant specifying the distance to an event source. This invariant can be incorporated into a virtual display, as appropriate, to enhance informativeness. As there is often a trade-off between speed and quantity of information delivery this remains a context dependent design decision.

Other differences in the way the two forms of energy can transmit information has to do with reflectivity. Light can be reflected around corners, but, under typical conditions, travels in straight lines (rays) and the information conveyed by a reflected ray is predominately specific to the reflecting surface. This contrasts with sound waves, which can reflect around corners, carrying information about the original sound source as well as the reflecting surfaces. When the distinctive speeds and reflective properties of light and sound are taken together, another informative distinction arises. Although both sound and light waves can be reflected, the speed of light makes those reflections undetectable to observers. Bregman (1990) has pointed out that this is not the case for slower travelling sound waves, and the generation of echoes can provide a great deal of information. In particular, the initial arrival of direct sounds (radiant) and subsequent reflected sounds (ambient) can be distinguished, providing separate information about the sound source and

the environmental properties (Gaver, 1993a). In comparison, radiant light alone provides little information for action (Heine & Guski, 1991), and the speed of light makes it impossible to distinguish between radiant and ambient light by arrival time. One additional, but critical point concerning reflectivity. Since sound can reflect around occlusions, it is able to provide information about events that are not optically specified at a given location. Again there is a context dependent trade-off here, this time between the informativeness of an echoic environment, and the degraded clarity of the signal.

Important differences also exist in the kinds of events that generate or influence each type of wave. While stationary objects are sufficient to reflect light rays, sound can only occur when there is movement to cause the vibration, and the propagation of the vibration itself through matter. Therefore auditory information is, by nature, about the movement of objects within the environment. As light requires no medium to propagate and sound does, any acoustic event carries information about the environment through which it travels. Every acoustic event can be expressed, at a basic level, by a four-stage model of vibration (Handel, 1989). A force (movement) initiates the vibrations. This force can be generated by numerous kinds of simple physical events such as impacts and dripping, that can be classified with respect to the states of matter involved (Gaver, 1993b). The initial vibrations are reinforced by resonations of physical bodies, and propagate through the environment. This implies that sounds will have acoustical qualities that are highly specific to the kinds of objects and forces that are involved in their generation. More generally, acoustical information is largely generated by objects themselves, as opposed to optical information, which is the result of reflections of light rays off of objects. Bregman (1990) expresses this distinction as a case of objects emitting sound, but reflecting light.

Typical acoustical environments are information-rich arrays consisting of direct and reflected sound waves generated by forces acting on various materials. The information in these arrays specifies the distance, location, and content of sound generating events. The characteristic reflectivity of sound waves means that in a typical situation, information from a source at any angle with respect to the listener can be picked-up. Accordingly, an obvious difference in the detection of optical and acoustical information is the degrees of freedom of movement of the eyes and ears. The eyes have a relatively limited area of information pick-up, but have multiple degrees of freedom, allowing efficiency of movement. In simple terms, the area of fixation can be altered by movement of the eyes themselves, movement of the eyes via the head, movement of the eyes via the body, or typically, some combination of the above. The ability to move the ears independently of the head is limited significantly in both range and utility, although the hands may be used to "focus" the pinnae. Thus, in order to gain a degree of movement with the ears, one must sacrifice a degree of manual control, an important constraint on interface design. As movement of the ears is predominately controlled by relatively gross changes in head and body position, active listening has the potential to be a more physically dynamic activity than looking. With respect to interface

design, both real and virtual sounds require head movements to differentiate positions in space. Therefore it is a key consideration to discourage precise auditory localization in situations where head motion is difficult, such as high g-force conditions.

There are clearly numerous differences in how location information can be delivered optically and acoustically. These features should be considering in assignment of data to be delivered to particular modalities. Recent work has utilized auditory and visual displays in conjunction with each other for the purposes of target localization and tracking. A number of studies have examined the effectiveness of 3-D auditory displays on audio/visual target detection performance. Although most of these investigators did not attempt to integrate auditory displays with other display systems, their work shows that spatially presented sound can be used to improve visual searches on a number of measures. McKinley, D'Angelo, Haas, Perrott, Nelson, Hettinger, and Brickman (1995) found lower workload scores in a simple detection task when participants were provided with localized audio, as compared with non-localized or no-sound conditions. In terms of performance, Perrott, Cisneros, McKinley, and D'Angelo (1996) found that the presentation of 3-D sound reduced target detection times, and that the utility of such a display increased as the target was presented further from the center of the visual field.

In an ecologically valid situation there would often be other visible objects in the field of view besides the target. Additional information about the targets position would be helpful, especially if that information was presented aurally, so as not to be masked by visual noise. This kind of situation was studied by Perrott, Sadralodabai, Saberi, and Strybel (1991). The authors found that in a cluttered field of view, 3-D sound was just as useful as an enhanced visual display of the target (such as in a Heads-Up-Display), but that at more distal positions the audio was a more effective aid than visual highlighting . Additionally, the reduction in detection times with the use of 3-D audio increased as a function of the degree of the visual load.

A recent paper by Bronkhorst, Veltman, and Breda (1996) is highly pertinent to the study being undertaken here. The authors looked at the effects of a 3-D auditory display system used separately and conjointly with a heads down threat location display. They found the best detection performance occurred in conditions which allowed the participants to utilize both displays, suggesting an advantage for simultaneous multimodal presentation. There were no overall differences in workload scores among the conditions in which at least one of the displays was present. The authors note that, to some degree, the visual and auditory displays were complimentary in that together they provided more information than either did alone, including the ability to disambiguate front-back confusions common to virtual auditory display systems. Accordingly, workload scores and search times were significantly greater for targets presented to the rear hemisphere, as compared with the front hemisphere. This may be explained by the effort of having to look down at a display and then back OTW to visually locate targets. The authors did not examine the effects of these displays on the efficiency of head and aircraft movements when

localizing targets, or other aspects of mission performance, as will be done in the present study. Also, unlike Bronkhorst et al. (1996), the present study will utilize a simulated head mounted display (HMD) pointer, analogous to the auditory display, to optically present information regarding target location.

The conjoint use of HMD and 3D sound for target localization has been effectively implemented. Flanagan, McAnally, Martin, Meehan, and Oldfield (1997), found that both visual and auditory cueing to targets outside of an HMD limited FOV were effective in reducing search times, with no performance difference between the modalities. Combining the displays resulted in a small reduction in search times, but no gain in search efficiency. Boucek, Orr, Williams, Montecalvo, Redden, Rolek, Cone (1996) obtained measures of threat acquisition and flight performance, as well as workload and situation awareness (SA) for simulated missions flown with baseline and advanced interface cockpits. In addition to the HMD and 3D sound, the advanced cockpit had a large screen tactical situation display (TSD), all of which contributed to the cockpit's overall greater effectiveness over baseline. Although the individual impact of each technology on performance increments was not assessed, the participants noted that the 3D audio was useful for providing initial threat azimuth location, and rated the HMD as contributing to target localization more than the TSD and audio displays.

As in the Boucek et al. study, a key research question the present study is whether the use of this display system will affect both performance and workload, not just for the target detection aspects of the task, but for the mission as a whole. The Boucek et al. study examined the advanced cockpit holistically, rather than looking at individual interface contributions. It is important to examine the role of individual multisensory displays, both alone and in tandem with other displays, to determine interface efficacy. For example, as information is presented to the auditory system there may not be a need to simultaneous present that information on a cockpit display or HUD. This would leave room for more detailed adaptive displays of pertinent systems data, and/or reduce visual clutter. Additionally, it is possible that reduced workload in the target localization aspects of the mission may positively affect performance during other navigational aspects of the mission.

Other improvements with the use of such a multi-modal system are not as readily observable, but may be important in improving performance and workload. The use of a head-slaved 3D audio system means that sounds will occur in virtual audio space with respect to the pilot's head position. If a sound comes from the far right and the pilot turns his head to visually focus on the target, then the sound will become centered, as would happen with a real sound-generating event. The overall effect of such a system is that targets appear with respect to the pilot, rather than the aircraft, creating a much more intuitive, directly manipulatable, and rapidly responding control loop with which to pick-up information. This might allow a pilot to attain a greater degree of situational awareness; i.e. knowing the location of targets with respect to his position. Using displays to make target localization into a more natural, automatic task follows

Vicente's and Rasmussen's (1992) recommendations for interface design. Specifically, keeping the task at a perceptual, rather than at a more cognitive demanding level (e.g. translating symbolic positions on a radar screen to actual locations in space), should result in faster, less effortful execution.

From a psychophysical perspective, this study can be viewed as an exploration of how people are capable of integrating information across perceptual systems. Wickens, Sandry, and Vidulich (1983) and Wickens (1992) provide evidence that cross-modal information exchange is generally more effective than intermodal delivery, especially in tasks that are unimodally demanding, such as visual scanning. Of particular relevance are situations where the perception of information concerning a given event covaries with modality. Since the auditory and visual displays in this project are computer generated, they will provide accurate, but not necessarily exact information about events in space. Accordingly, the information provided simultaneously to the auditory and visual channels may not correspond. O'Leary and Rhodes (1984) demonstrated that both asymmetrical visual and auditory stimuli could have cross-modal effects on the perception of other stimuli, that is, information received by one sense could modify the perception of information to another sense. It will be interesting to examine the conditions that determine which modality is more influential on a pilot's perception of target location. These kind of findings can go back into the design cycle to help determine the optimal modality(ies) for information display over mission situations.

To better understand these complex relations between multimodal displays and operator behavior, analogous auditory and visual spatial information displays were developed to be used in a demanding flight/target localization task. Assessments of target detection and flight performance, along with workload measures, will be collected across a number of different auditory and visual display integrations.

Method

•Summary

Pilots flying a simulated target lock-on and course navigation mission will utilize a head-coupled on screen cursor to designate visually sighted targets. Target position displays consist of head-coupled visual and auditory pointers that present angular distance and range information. Tracking efficiency and percentage of designated targets, among other measures, will be used to assess performance under varying single and integrated multisensory display conditions. Flight path information will be presented through a head-up and heads down displays, and navigational error will be a function of spatial and temporal course deviations. Workload for the subtasks will be

assessed using a combination of the SWORD and TLX surveys.

•Participants

Approximately twelve military pilots will be recruited for this experiment. Participants will meet a minimum criteria of flight hours, and have normal or corrected-to-normal vision and hearing. Participants will be compensated and take part in training and experimental trials over the course of two consecutive days.

•Experimental Design

A completely within subjects design will be utilized in order to maximize the number of observations, and to examine individual differences in performance and workload across conditions. The order of the seven different display configurations (see Table 1) will be balanced across subjects.

•Apparatus

This study will be conducted in the Synthesized Immersion Research Environment (SIRE) Laboratory. The SIRE uses multiple, high-definition Barco projectors to generate a 150 degree horizontal x 70 degree vertical out the window field of view. Graphics are generated by a Silicon Graphics Onyx computer and presented on a semi-circular dome. Pilots will be seated in a model open canopy F-16 cockpit (see Figure 1). The cockpit contains a control stick with trigger, throttle control, and head down avionics displays which present basic flight instrumentation. A head-up-display (HUD) resides on top of the cockpit to correspond with the pilots lower boresight field of view. The HUD will graphically present information on course deviation and corrective action in addition to airspeed, altitude, and other basic flight variables.

Figure 1. *Partial view of pilot seated in cockpit with OTW view.*

The seven interface configurations consist of combinations of no pointer, a visual pointer, and an auditory pointer (see Table 1). The visual pointer is a head-slaved dome projection

(simulating a head mounted display) that provides target direction and distance information with respect to the pilot's head position. An arrow, whose origin is at the center of the pilot's head controlled target lock-on cursor, will extend from the center of the field of view to a scaled distance representative of the angular distance to the threat (see Figure 2). The completeness of circle surrounding the arrow will indicate range to target. For example, half of maximum range would be indicated by a half-circle.

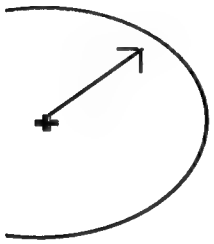


Figure 2. *Visual angular distance and range display for a target at approximately 40 degrees visual angle and half maximum sensor distance.*

The auditory pointer utilizes headphone presented 3-D sound to display target position, also with respect to head position. Range will be indicated by the relative intensity of a broadband noise pulse. As in Bronkhorst et al. (1996), measurements of listeners' head-specific acoustical features will be used to generate individualized head-related transfer functions (HRTFs). This enables the Convolvotron sound computer to present sources in virtual space, customized to the listener, that sound significantly more realistic than if a generic transfer function had been used.

In the adaptive conditions, an adaptive algorithm will switch display configurations whenever the nearest target crosses a FOV boundary. It is based on Gibson's values for a field of view that extends approximately 80 degrees laterally and 60 degrees vertically (Perrott, Sadralodabai, Saberi, & Strybel, 1991). For example, in configuration seven (Table 1) a target out of the FOV will generate an auditory display. When the target enters the FOV boundary the auditory display will be supplemented with visual display.

•Procedures

Pilots will be trained to follow a low-altitude, directed flight path over a database of New Mexico terrain, while searching for and locking on targets with the use of a given display configuration. Pilots will designate a target by placing a head-coupled target-box over the threat and depressing a control stick trigger. A positive designation will be signified by a brief change in target-box color corresponding with the presentation of a lock-on tone. Numerous flight paths will be generated following a sum of sines algorithm to allow for equal complexity. All pilots will fly the same set of paths, balanced for order and display condition. During the data collection phase participants will fly blocks of 5 trials per display condition, lasting approximately 1/2 hour per block.

Following each block the workload measures will be taken, and following the final block a debriefing questionnaire will be administered.

Table 1.
Field-of-View Dependent Interface Configurations.

		INTERFACE CONFIGURATION	CHANNEL FOR TARGETS IN FOV	CHANNEL FOR TARGETS BEYOND FOV
NON-ADAPTIVE	1	Non-cueing	-	-
	2	Full time Visual	Visual	Visual
	3	Full time Auditory	Auditory	Auditory
	4	Full time Visual and Auditory	Visual and Auditory	Visual and Auditory
ADAPTIVE	5	Visual or Auditory	Visual	Auditory
	6	Visual or Visual and Auditory	Visual	Visual and Auditory
	7	Visual and Auditory or Auditory	Visual and Auditory	Auditory

Expected Results

The anticipated benefits of the different interface configurations will be evaluated by multiple hypotheses. Please refer to Table 1 for interface configuration descriptions. Individual hypotheses will be evaluated using analyses of variance with adjustment for repeated measures.

A)Target cueing enhances performance and manages workload: 1 vs 2-7.

It is expected that target lock-on rates will be improved dramatically when pilots are provided with target position and range information versus when they must search for it unaided. Additionally, the reduction of the searching task should result in less workload associated with target detection.

B)Visual target detection is enhanced by 3D sound, especially for targets beyond FOV: 2 vs 4-7, but also for targets in the central FOV: 4,7 vs 2,5,6.

Based on several of the studies cited earlier, the presence of correlated 3D sound should improves target detection rates and times for all target locations, with improvement increasing as a function of distance from the center of the field of view.

C)Certain multisensory interfaces improve performance/reduce workload in comparison to single modality interfaces: 2,3 vs 4-7.

Presenting information through multiple channels should make information more salient through information richness and user self-selection of the of the more efficient channel, This should reduce multiple task workload.

D)Appropriate (FOV dependent) modality switching of target position information give as good performance/workload results as full-time multimodal display: 4 vs 5-7.

Pilots may switch between the present channels based on the utility of the channels for the situation. Appropriate adaptive channel switching could then be expected to be as useful as full-time multichannel display.

E)Combining multimodal adaptive interface with fixed interface enhances performance over single modality presentation: 2,3,5 vs. 6,7.

Pilots will be provided with appropriate information display that is modulated for current field of regard. This should result in more rapid target acquisition than a fixed modality display which may have greater perceptual and/or cognitive demands in certain situations.

F)Auditory integration will enable tracking while managing visual attention to flight displays, implicitly reducing navigational course errors. 2 vs 3-7.

With localization aiding displays, more attention may be devoted to navigational tasks, lowering flight course deviations.

There will be several significant payoffs from this research including the generation of paradigms for the assignment of threat information to auditory and visual displays. These could be applied to displaying other kinds of information, as well as to other multisensory display systems. Moreover, a demonstration of an operationally effective mixed modality display may "free-up" the fixed roles of visual displays and thereby allow for the development of dynamic, adaptive technology. This work should help direct the design and focus of further research into adaptive and multisensory information display and control. By helping to define what kind of situation-dependent interface changes produce effective changes in performance, we are better able to

evaluate the variables that may determine when systems should adapt. For example, multimodal information portrayal may be combined with other research, such as physiological workload assessment measures, to create a real time situational and pilot-state adaptive algorithm. These kinds of products will promote air superiority by giving pilots more direct perception and control of their aircraft and sub-systems (Haas, 1996).

References

Bennett, K.B. (1997). *Dynamically adaptive interfaces: A preliminary investigation - Final report for: Summer Research Extension Program, Armstrong Laboratory.* Unpublished Manuscript.

Boucek,G.S., Orr,H.A., Williams,R.D., Montecalvo,A.J., Redden,M.C., Rolek,E.P., and Cone, S.M. (1996). *Integrated mission precision attack cockpit technology (IMPACT). Phase 2. Cueing benefits of large tactical situation displays, helmet-mounted displays, and directional audio.* Air Force Materiel Command WL-TR96-3076.

Bregman, A.S. (1990). *Auditory Scene Analysis.* Cambridge, MA: MIT Press.

Bronkhorst, A.W., Veltman, J.A., & Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human Factors*, 38, 23-33.

Cutting, J. E. (1986). *Perception with an Eye for Motion.* Cambridge, Mass.: MIT Press.

Flanagan, P., McAnally,K.I., Martin,R.L., Meehan,J.W., and Oldfield, S.R. (1997). *Search time and efficiency for aurally and visually guided visual search in a virtual environment.* Unpublished manuscript, Deakin University.

Gaver, W.W. (1993a). What in the world do we hear?: An ecological approach to auditory even perception. *Ecological Psychology*, 5, 1-29.

Gaver, W.W. (1993b). How do we hear in the world?: Explorations in ecological acoustics. *Ecological Psychology*, 5, 285-314.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston: Houghton-Mifflin.

Haas, M. (1996). *New World Vista Proposal, Project 11: Human Machine Interface, Research Area: Information Management and Display.* USAF.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Heine, W-D., & Guski, R. (1991). Listening: The perception of auditory events? *Ecological Psychology*, 3, 263-275.

Lee, D.N. (1990.) Getting around with light or sound. In R. Warren & A.H. Wertheim (Eds.), *Perception and control of self-motion.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McKinley, R.L., D'Angelo, W.R., Haas, M.W., Perrott, D.R., Nelson, W.T., Hettinger, L.J., & Brickman, B.J. (1995). *An initial study of the effects of 3-dimensional auditory cueing on visual target detection.* Poster Session at Human Factors and Ergonomics Society 39th Annual Meeting. San Diego, CA. October.

Moroney, B. (1997). *The role of multimodal adaptive interfaces in providing flight path guidance. Final report for: Graduate Summer Research Program, Armstrong Laboratory.* Unpublished Manuscript.

O'Leary, A. & Rhodes, G. (1984). Cross-modal effects on visual and auditory object perception. *Perception & Psychophysics*, 35, 565-569.

Perrott, D.R., Cisneros, J., McKinley, R.L. , & D'Angelo, W.R. (1996). Aurally aided visual search under virtual and free field listening conditions. *Human Factors*, 38, 702-715.

Perrott, D.R., Sadralodabai, T., Saberi, K., & Strybel, T.Z. (1991). Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target. *Human Factors*, 33, 389-400.

Stein, B.E., Meredith, M.A., and Wallace, M.T. (1994). Development and neural basis of multisensory integration. In D.J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp.81-106). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Vicente, K.J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 589-06.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance.* New York: Harper Collins.

Wickens, C.D., Sandry, D.L. and Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, 25, 227-248.

A PRELIMINARY EXAMINATION OF ECL ACTIVITY GEARED TOWARD A CD^{+2} SENSOR

Paul J. Taverna
Graduate Student
Department of Biomedical Engineering

Tulane University
500 Lindy Boggs
New Orleans, LA 70118

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1997

A PRELIMINARY EXAMINATION OF ECL ACTIVITY GEARED TOWARD A CD^{+2} SENSOR

Paul J. Taverna
Graduate Student
Department of Biomedical Engineering
Tulane University

Abstract

Pollution is a growing concern of the United States Air Force. As a result, there is a pressing need for accurate methods to determine metallic concentrations in polluted water, sediment, and effluent from certain metallic processes. The technique of electrochemiluminescence (ECL) has previously shown promise in low concentration assays for organic compounds, but not extensively explored for specific use as a metal sensor. It was hypothesized that ECL could be exploited to quantify metallic concentrations using reagents that show specific enhancement for an ion of interest. Twelve organic compounds were examined for ECL activity in the presence of tripropylamine (TPA) and 34 metal ions. Of the 34 ions screened, cadmium, showed enhanced ECL activity when combined with 1,10-phenanthroline, terpyridine, and bipyridine. The ligand parameters were optimized for each ligand to encompass the effect of organic ligand concentration, TPA concentration, TPA pH, time upon signal, and mixing order. A cadmium calibration curve was compiled to examine the degree of linearity after ligand parameters were optimized for maximum signal generation for each organic compound. Interference of additional ions with cadmium were also examined and found to be problematical with ten fold amounts of interfering ions to cadmium ions, however caused smaller interference at 1 ppm levels and thought to be due to a competition effect with cadmium ions for interaction with organic ligands. The best limit of detection obtained with the optimized protocol of 1,10 -phenanthroline was found to be 9 ppb, 1 ppb below the accepted 10 ppb EPA level for cadmium in drinking water.

A PRELIMINARY EXAMINATION OF ECL ACTIVITY GEARED TOWARD A CD^{+2} SENSOR

Paul J. Taverna

Introduction

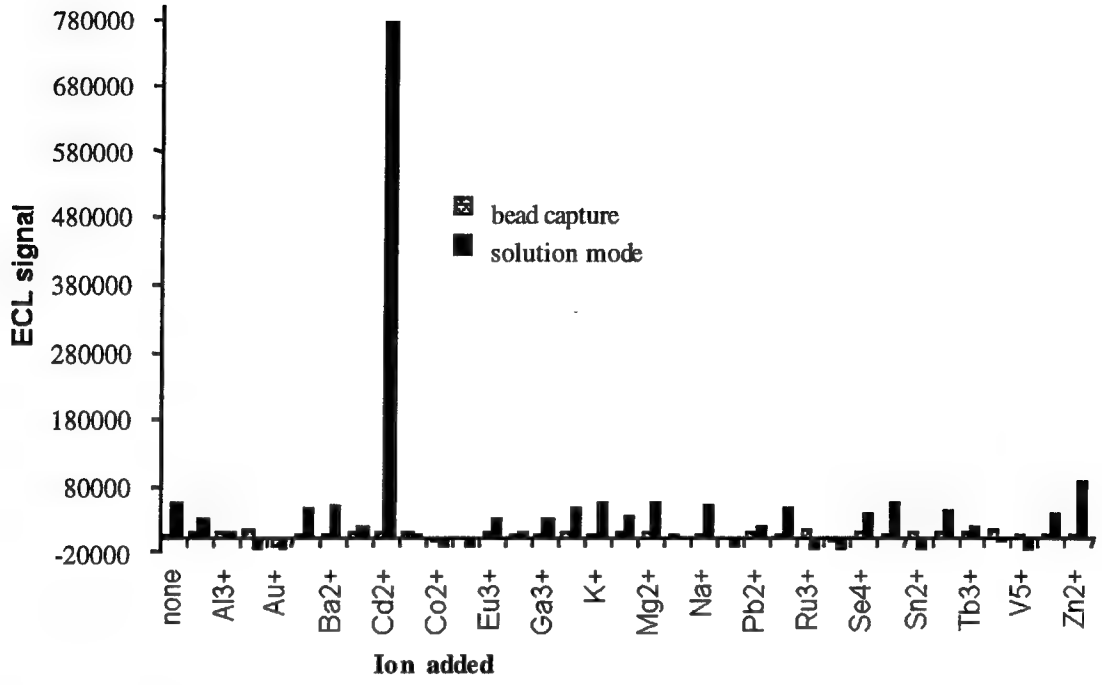
Electrochemiluminescence (ECL) is a phenomena that has been utilized in several types of assay techniques[1-3]. For example in a well characterized ECL reaction, a tripropylamine (TPA) molecule is complexed with a ruthenium (II) trisbipyridine ($Ru(bpy)_3^{2+}$) molecule once: 1) $Ru(bpy)_3^{2+}$ has been oxidized electrochemically to $Ru(bpy)_3^{3+}$ and 2) TPA oxidized to a free radical (TPA^*), both at the electrode surface. The two oxidized compounds join to form a short-lived complex that relaxes to TPA byproducts and $Ru(bpy)_3^{2+}$, with a photon released at a wavelength of 620 nm. The $Ru(bpy)_3^{2+}$ is recyclable in ECL reactions, however the TPA byproducts are not[4]. It has been suggested that for a sample of known volume the concentration of unknown metal ions complexed with the $Ru(bpy)_3^{2+}$ can be determined by correlating the sample ECL measured to ECL measured of known concentrations standards. It has also been suggested that other organic compounds could replace $Ru(bpy)_3^{2+}$ as the metal ligand.

Methodology

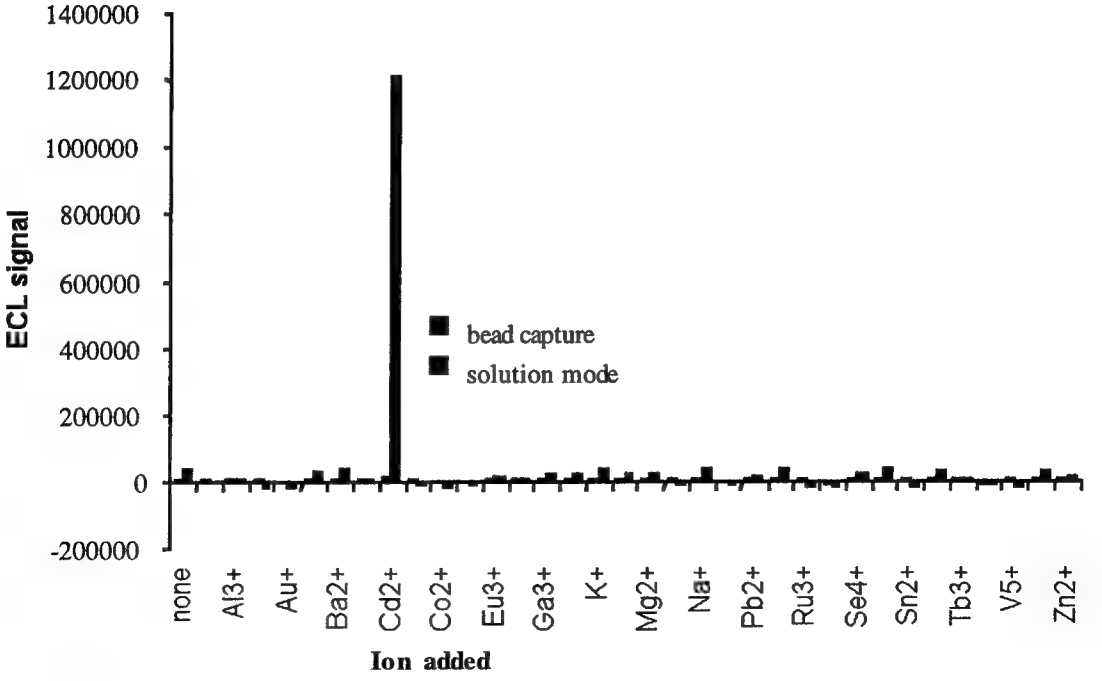
Ten organic compounds containing nitrogen in their aromatic structures were examined for ECL when combined with 34 metal ions. The preliminary screening was performed after the addition of 700µl 0.2 M TPA, 200 µl organic ligand, and 100 µl ion. The samples were allowed to sit for 15 minutes and assayed using an Origen® Analyzer, Igen Corporation, MD.

The organic compounds examined in the 34 metal ion screening were: Phthalazine, quinazoline, quinoxaline, 1,10-phenanthroline, 4,7-phenanthroline, 2,2':6',2"-terpyridine, phenazine, 2-3-bis(2 pyridyl)pyrazine, 2,3-diphenyl-1,4-diazaspiro[4,5]deca-1,3-diene, 2,3-diaminonapthaline, 2,2'-bipyridine, and benzo[c]cinnoline. Each compound was dissolved in methanol at a concentration of 1mg/ml. The elements added to each organic compound were 18 MΩ H_20, Ag, Al, As, Au, B, Ba, Ca, Cd, Cr, Co, Cu, Eu, Fe, Ga, Hg, K, Li, Mg, Mn, Na, Ni, Pb, Re, Ru, Sb, Se, Si, Sn, Sr, Tb, Tl, V, W, and Zn. Each element added was at a concentration of 1mg/ml. For further details on the ion screening, see "Novel Electrochemiluminescence Reactions and Instrumentation," Anthony Andrews, AFOSR, Washington, DC, Armstrong Laboratory/Environics Directorate, Tyndall AFB, FL, 1997.

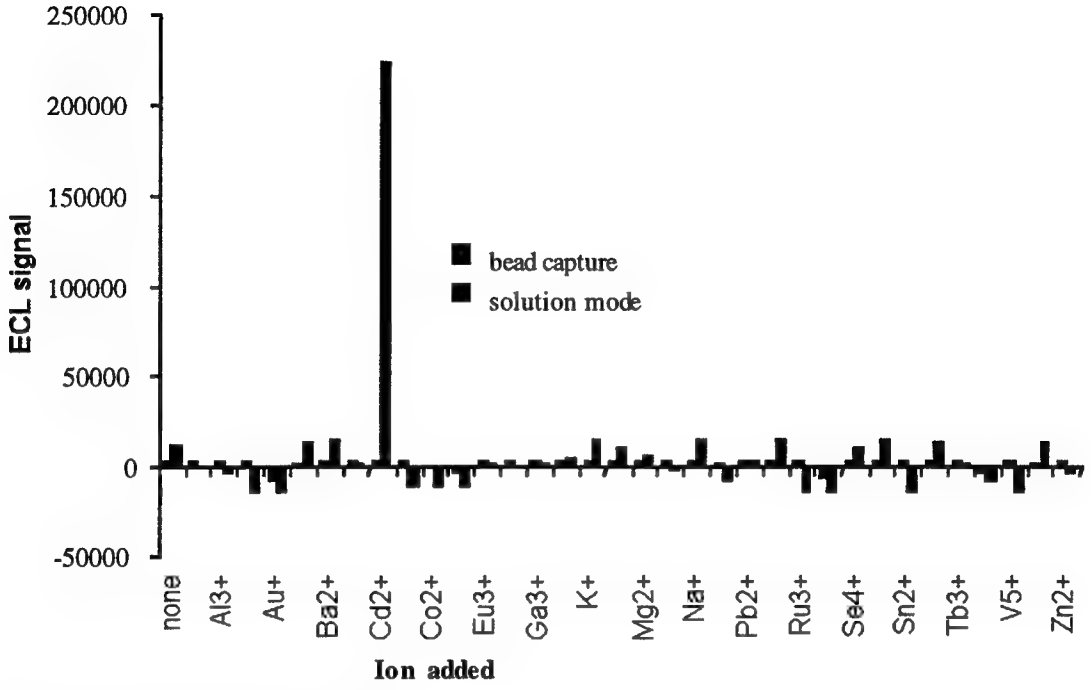
Three of the twelve organic ligand compounds examined showed a large signal when assayed w ith divalent cadmium (Cd^{+2}) and TPA. The compounds were 1,10-phenanthroline, terpyridine, and bipyridine. The results (with selected ions listed) are shown below.



34 Metal Ion Screen with 1,10-phenantholine.



34 Metal Ion Screen with Terpyridine.



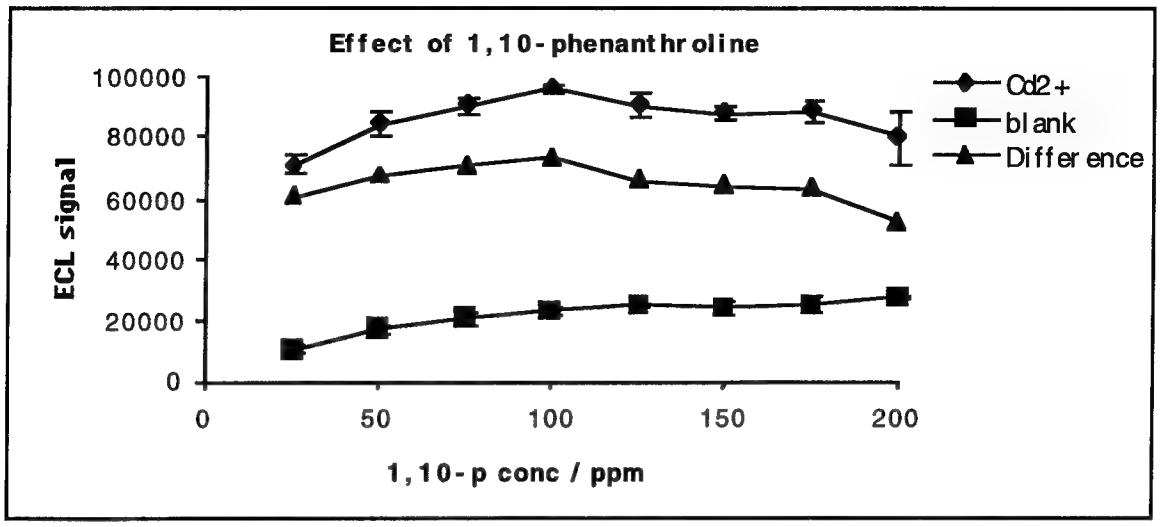
34 Metal Ion Screen with Bipyridine.

1,10-phenanthroline

Further examinations of 1,10-phenanthroline (1,10-P) were performed to maximize ECL attributed to cadmium and minimize ECL due to reagents in a sample. With each ECL measurement for the parameters examined below, two samples were prepared. One contained cadmium and the other contained DI water in place of cadmium to act as a blank. The water blank was subtracted from the cadmium containing sample after measurement and the difference recorded, adjusting for ECL inherent to reagents used.

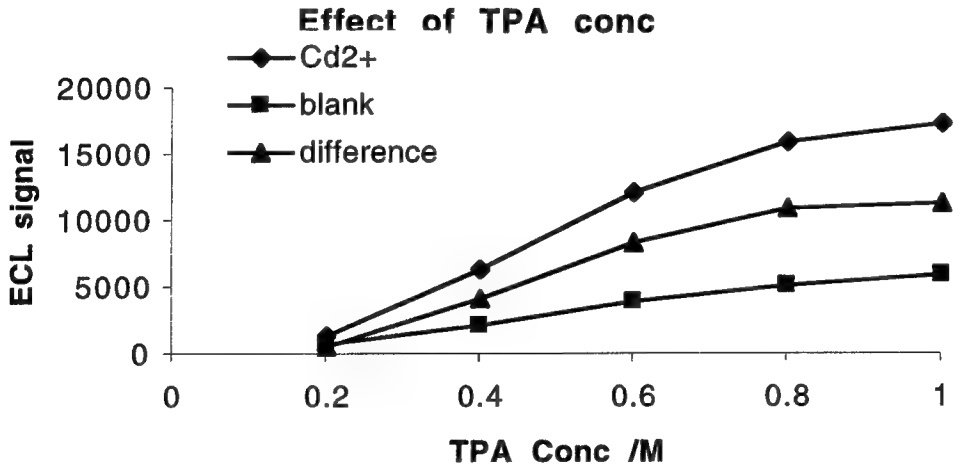
Effect of 1,10-P concentration

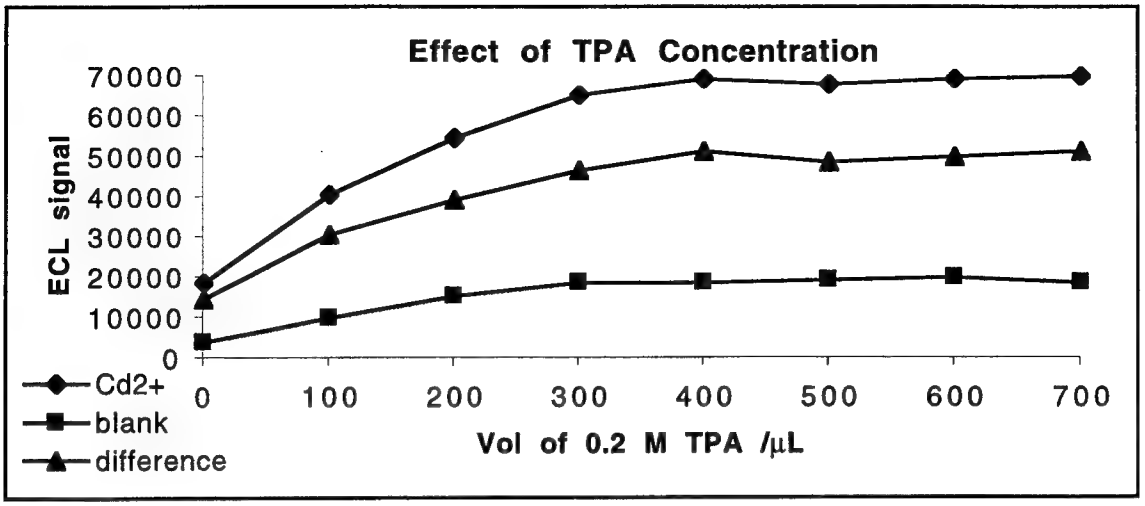
1,10-P concentration from 0-200 ppm was first examined using stock 1mg/ml 1,10-P in dilutions of 0-200 ppm and second, diluted stock of 0.1 mg/ml for 0-20 ppm examinations. A visible peak at 20 ppm 1,10-P was observed and increasing 1,10-P concentration above 20 ppm showed a decrease in ECL output. These findings suggest 200µl, 0.1 mg/ml 1,10-P is ideal. The final concentration in each 1 ml sample is 20 ppm.



Effect of TPA Concentration

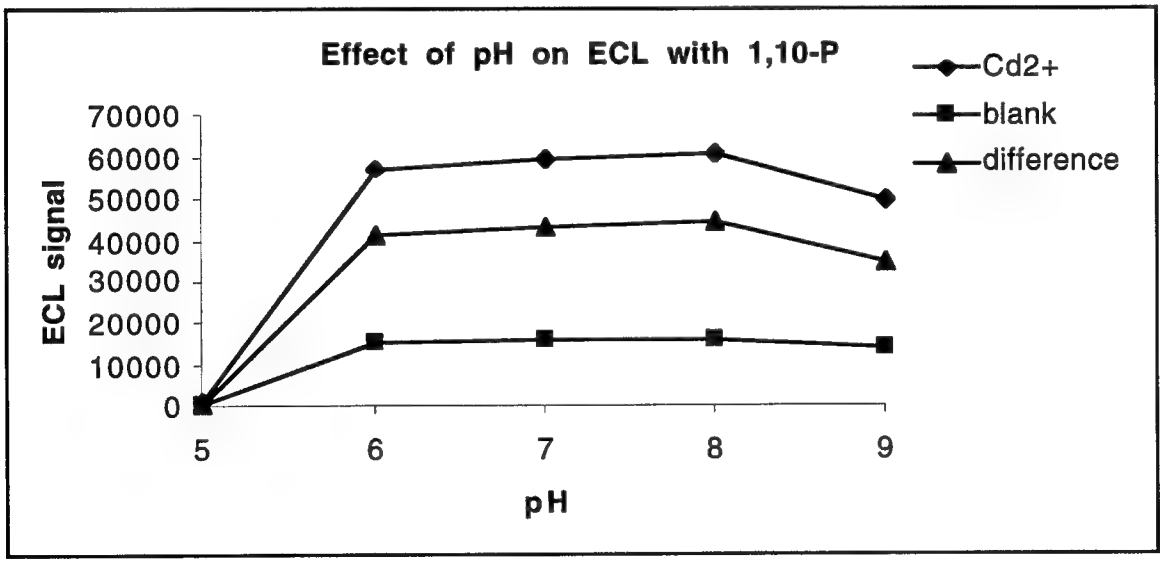
The volume of 0.2 M TPA per sample was varied from 0 to 700 µl. The ECL output was shown to be linearly increasing after 200 µl. The TPA concentration from 0.2 to 1 M was also examined. The result appeared to be a linearly increasing ECL output with increasing TPA concentration. However, the magnitude of ECL output at concentrations of 0.2-1.0 M was much lower than those of 0.2 M examinations. Therefore to maximize ECL, 700 µl 0.2 M TPA was used in subsequent analysis.





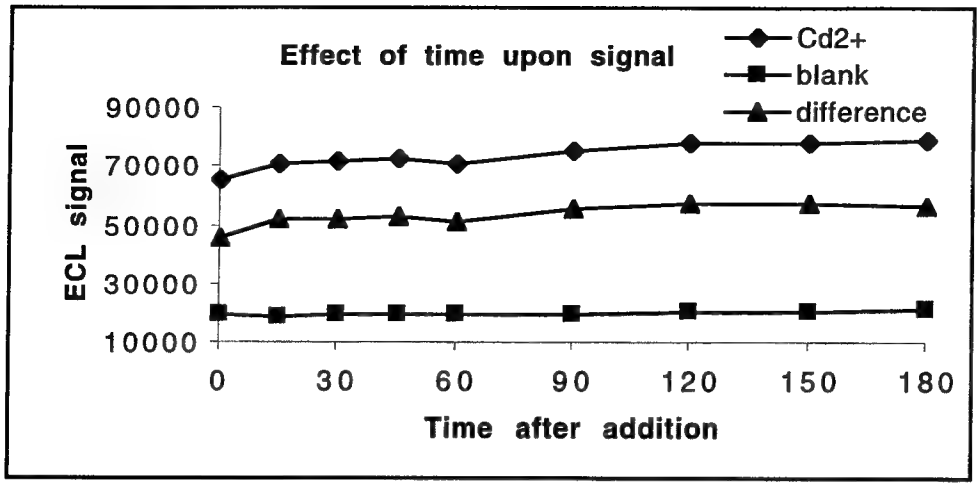
Effect of TPA pH

The effect of TPA pH upon ECL was examined at pH values of 5, 6, 7, 8, and 9 using 200µl 0.1 mg/ml 1,10-P, 700 µl 0.2 M TPA, and 100 µl 10 ppm Cd^{+2}. A peak in ECL was noticed between 7 and 8. Additional pH values were tested at 7, 7.2, 7.4, 7.6, 7.8, 8, 8.2. Further examinations included pH values of 8, 8.2, 8.4, 8.6, and 8.8 . From the data collected below, TPA at a pH of 8.2 appeared to maximize the ECL signal for cadmium.



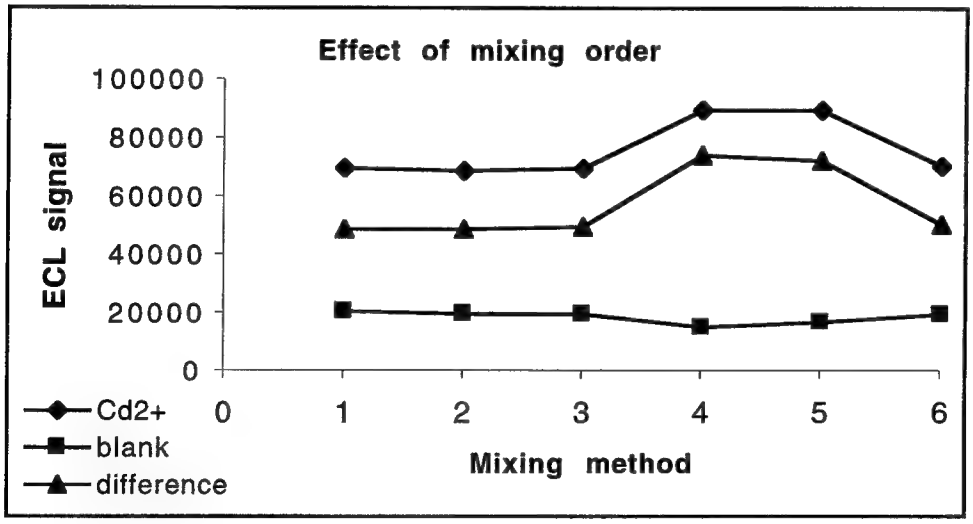
Effect of Time Upon Signal

Samples were mixed in the previously described order and allowed to interact for 0, 15, 30, 45, 60, 90, 120, 150, and 180 minutes. Allowing the reagents to mix for excess of 15 minutes showed little improvement on the ECL signal, as seen below.



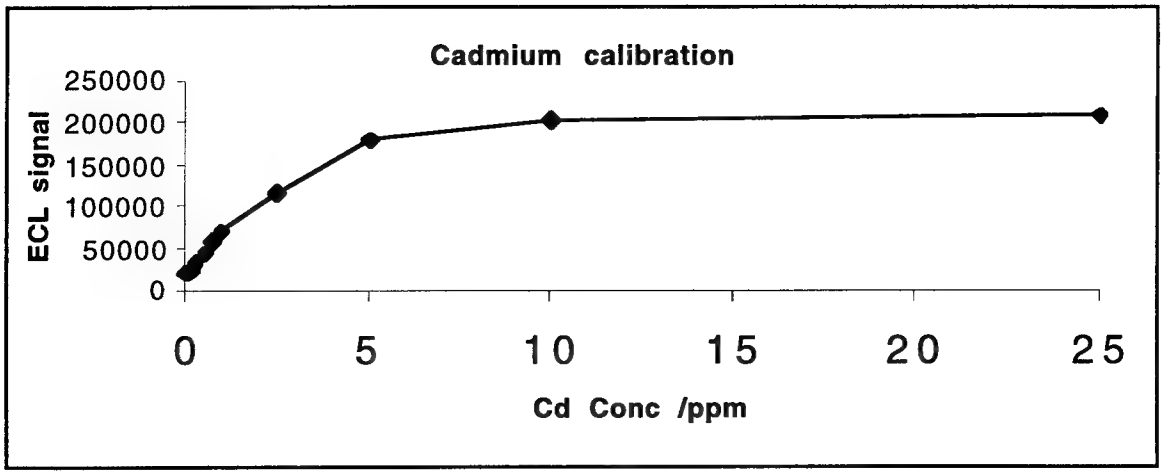
Effect of Mixing Order

The samples were examined for any ECL enhancement due to mixing order. The orders were as follows: 1)ion, 1,10-P, TPA, 2) ion, TPA, 1,10-P, 3)1,10-P, ion, TPA, 4) 1,10-P, TPA, ion, 5) TPA, 1,10-P, ion, 6) TPA, ion, 1,10-P. The concentrations and amounts of reagents added were determined in the parameters investigated before mixing order. The output of the experiment showed mixing method 4 to be highest.



Cadmium Calibration

A cadmium calibration curve compiled with measurements of 0-25 ppm was performed utilizing the optimum parameters of 1,10-P concentration, TPA concentration, TPA pH, time upon signal, and mixing order as described above. The linear range of the calibration curve was from 0-1 ppm, with an r^2 = 0.9994 . A repeat of the curve produced an r^2 = 0.9963.



Interference for Added Ions

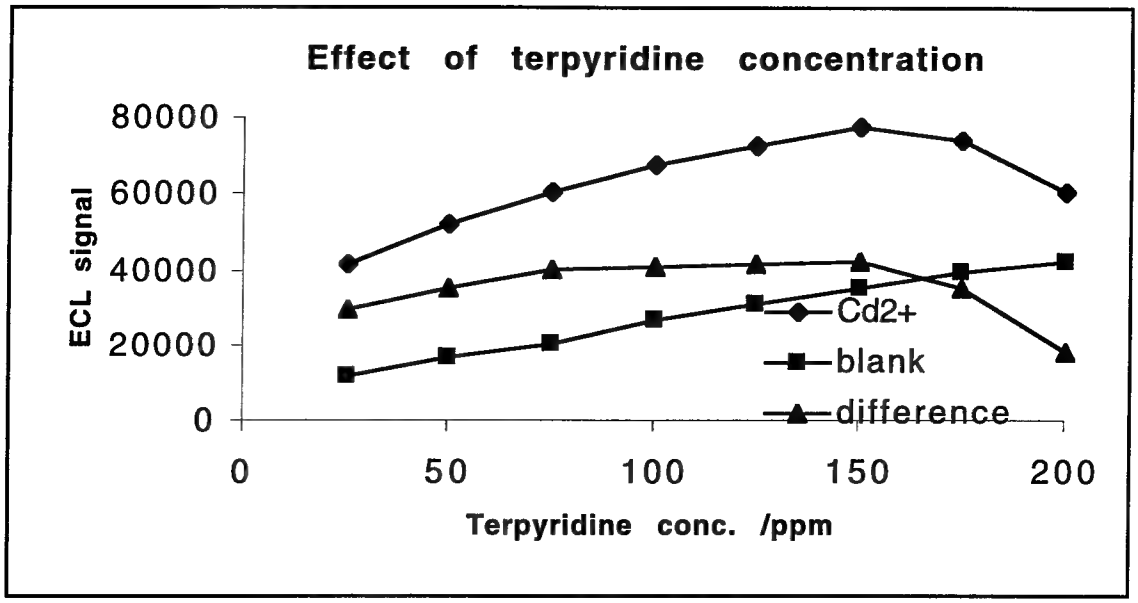
Using the optimized quantities and concentrations of the parameters determined, an ion interference run was performed. The ions of potential interference (100 µl of 1000 ppm standards) were added to 200 µl of 1,10-P, 600 µl TPA, and 100 µl of Cd^{+2}. The potential interfering ions were DI H_2O (control), tap H_2O, Al, Ba, B, Ca, Co, Cu, Fe, Mg, Ni, K, Se, Na, Sn, V, and Zn. Final concentrations in the cell were 100 ppm for the ions added, 1 ppm for Cd^{+2}. Mostly diatomic ions showed a large reduction in the Cd^{+2} control signal. Ions that caused a reduction in signal included Al (17.5% of control ECL), Ca (3.3%), Co (6.7), Fe (40%), Mg (6.4%), Ni (7.2%), Sn (1.2%), V (0.4%), Zn (3.9%). Repeating the above interference test with 10 ppm final ion concentrations (100 µl of 100 ppm ion into 1 ml test cell) showed signal inhibition from Co (10.0%), Cu (26.8%), Ni (20.2%). At 1 ppm (100 µl of 10 ppm standard) the smallest inhibition of the Cd signal was Co at 79.9% of control ECL measurement..

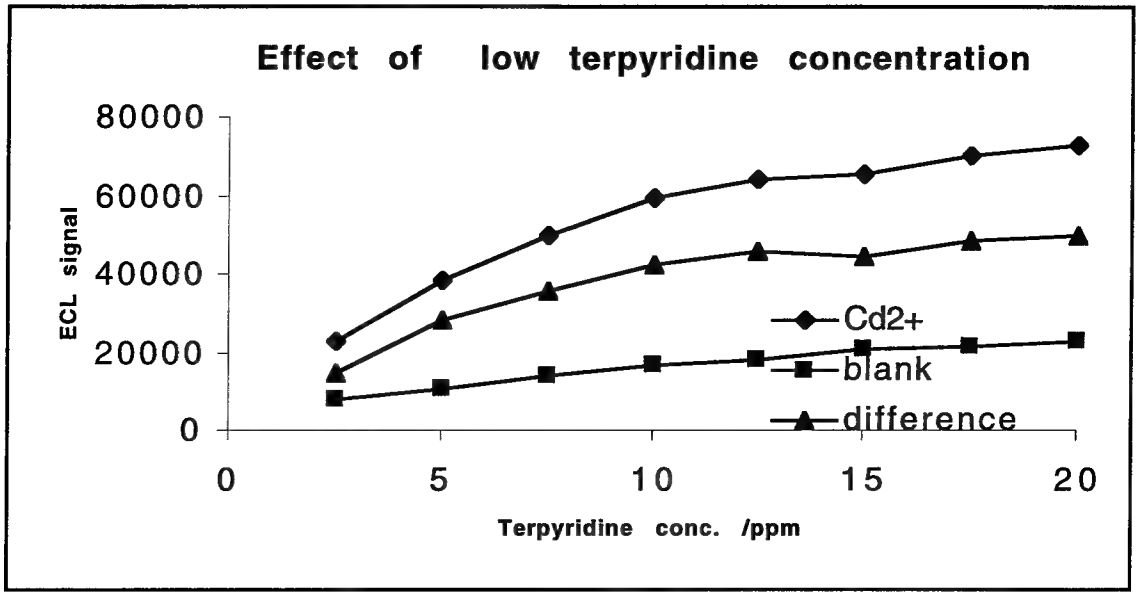
Terpyridine

The examinations of 1,10-P parameters were repeated for terpyridine. As with each ECL measurement of 1,10-P parameters, two samples types were prepared. Three tubes contained cadmium and the other three tubes contained DI water as a blank. The water blank was subtracted from the cadmium containing sample after measurement and the difference was recorded.

Effect of terpyridine concentration

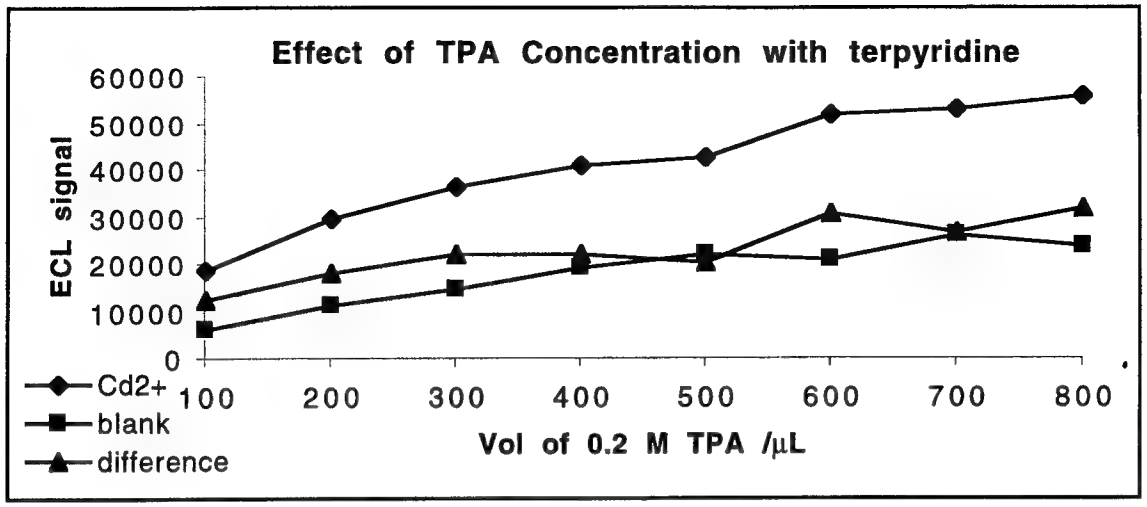
Terpyridine concentrations from 0-200 ppm were examined using stock solution of 1 mg/ml, and 0-20 ppm using a dilute solution of 0.1 mg/ml. The ECL values obtained at terpyridine concentration from 75 to 100 ppm were similar in value as shown below. This find suggested 100 µl 0.1 mg/ml terpyridine is ideal in terms of addition simplicity when using an automatic pipetting device. The final concentration in the 1 ml sample is 10 ppm.





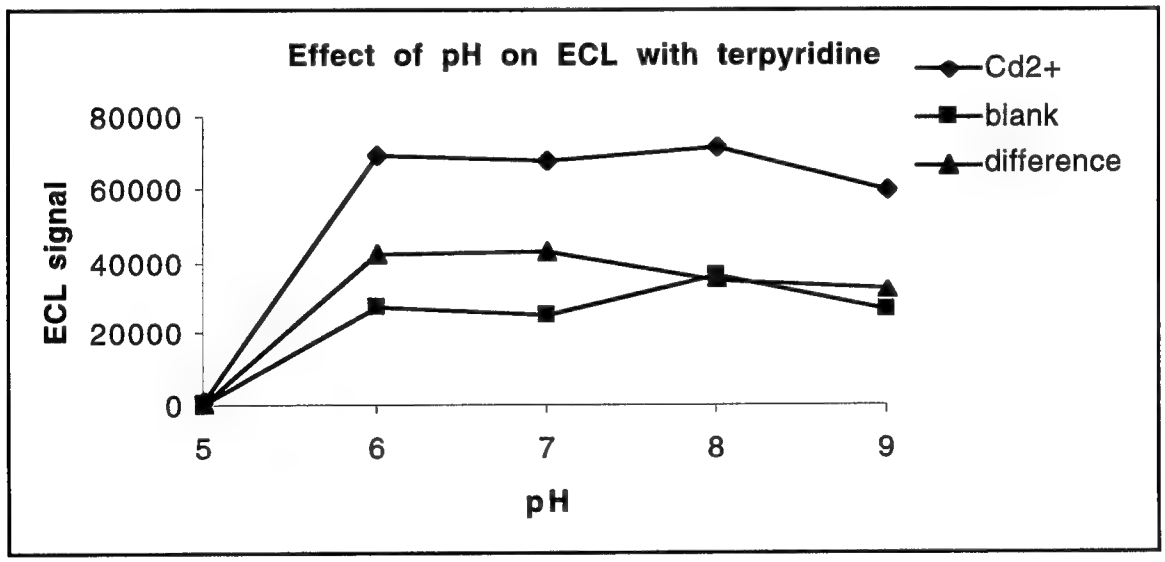
Effect of TPA Concentration

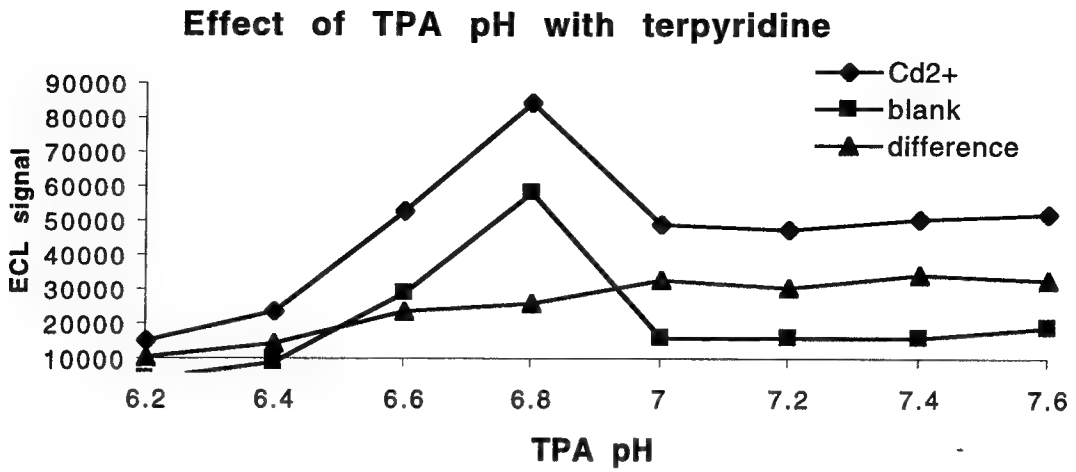
The volume of 0.2 M TPA was varied from 0 to 800 µl per each run. The ECL output was shown to be linearly increasing with increasing TPA volume. No advantage was found going below 800 µl per sample.



Effect of TPA pH

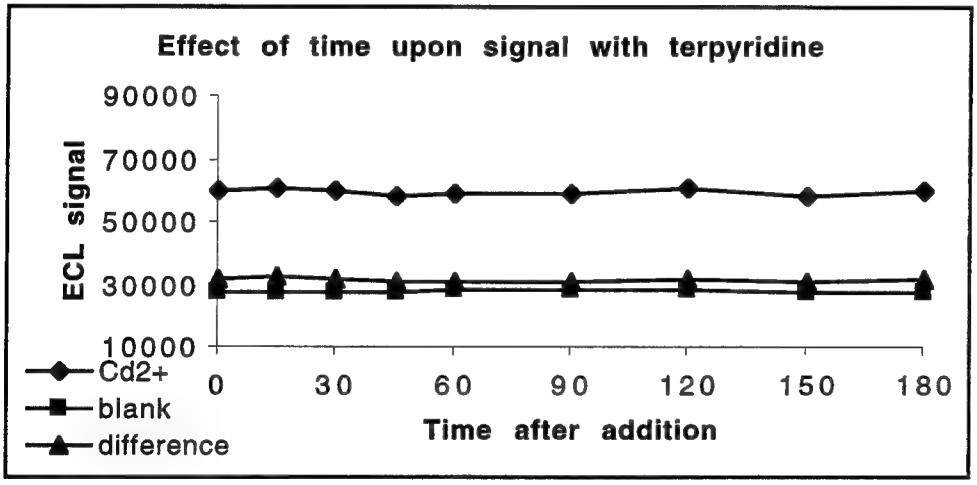
The effect of TPA pH was examined at pH values of 5, 6, 7, 8, and 9 using 100 µl 10 ppm Cd^{+2}, 100 µl 0.1 mg/ml terpyridine, and 800 µl 0.2 M TPA. A peak in ECL was noticed between 6 and 8. Further pH values were examined at 6.2, 6.4, 6.6, 6.8, 7, 7.2, 7.4, and 7.6. From the data collected, TPA at a pH of 7.5 maximized the ECL signal when exposed to cadmium.





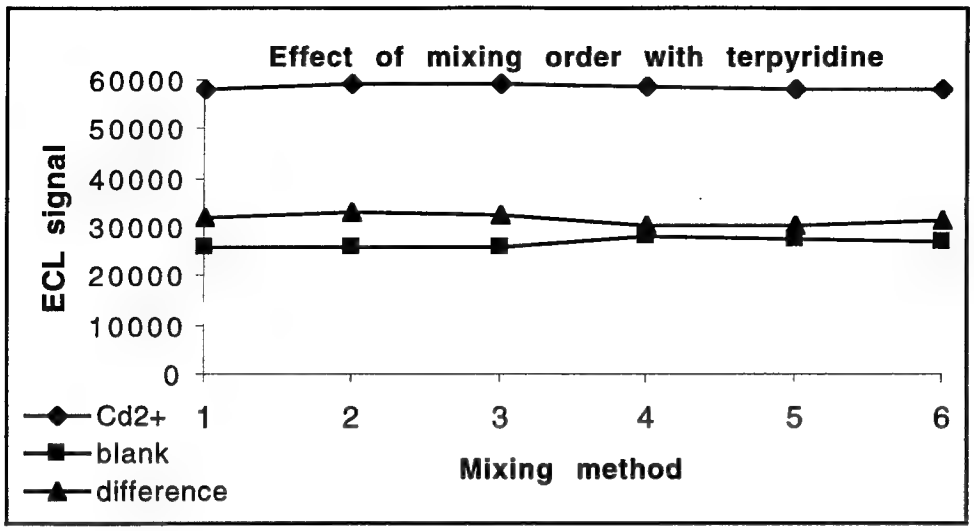
Effect of Time Upon Signal

Samples were mixed in the previously described order and allowed to interact for 0, 15, 30, 45, 60, 90, 120, 150, and 180 minutes. Allowing the reagents to mix for excess of 15 minutes showed little improvement on the ECL signal.



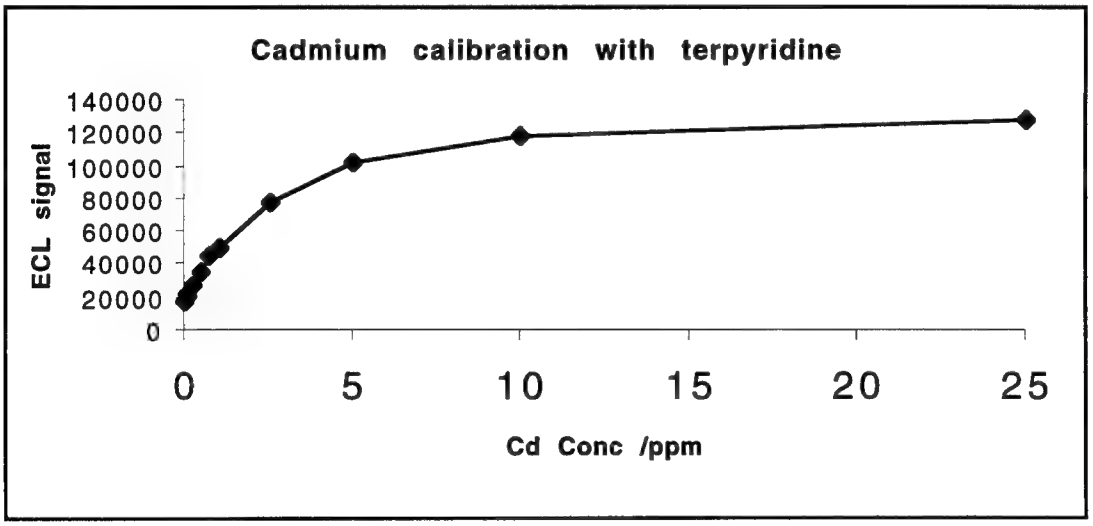
Effect of Mixing Order

The samples were examined for any ECL enhancement due to mixing order. The orders were as follows: 1)ion, terpyridine, TPA, 2) ion, TPA, terpyridine, 3) terpyridine, ion, TPA, 4) terpyridine, TPA, ion, 5) TPA, terpyridine, ion, 6) TPA, ion, terpyridine. The concentrations and amounts of reagents added were the ones determined in the examinations previously described. The output of the experiment showed mixing method 2 to be highest of those examined.



Cadmium Calibration

A cadmium calibration curve compiled with concentrations of 0-25 ppm was constructed utilizing the optimized parameters of terpyridine concentration, TPA concentration, TPA pH, time upon signal, and mixing order as found above. The linear range of the calibration curve was from 0-1 ppm, with an r^2 = 0.9922 .



Interference for Added Ions

Using the optimum reagent quantities and concentrations for the parameters investigated above, an ion interference experiment was performed. The ions of potential interference (100 µl of 1000 ppm standards) were added to 100 µl of terpyridine, 700 µl TPA, and 100 µl of Cd^{+2}. The ions were DI H_2O (control), tap H_2O, Al, Ba, B, Ca, Co, Cu, Fe, Mg, Ni, K, Se, Na, Sn, V, and Zn. Final concentration in the cells were 100 ppm for the ions added, 1 ppm for Cd^{+2}. As the examination with 1, 10-P, mostly

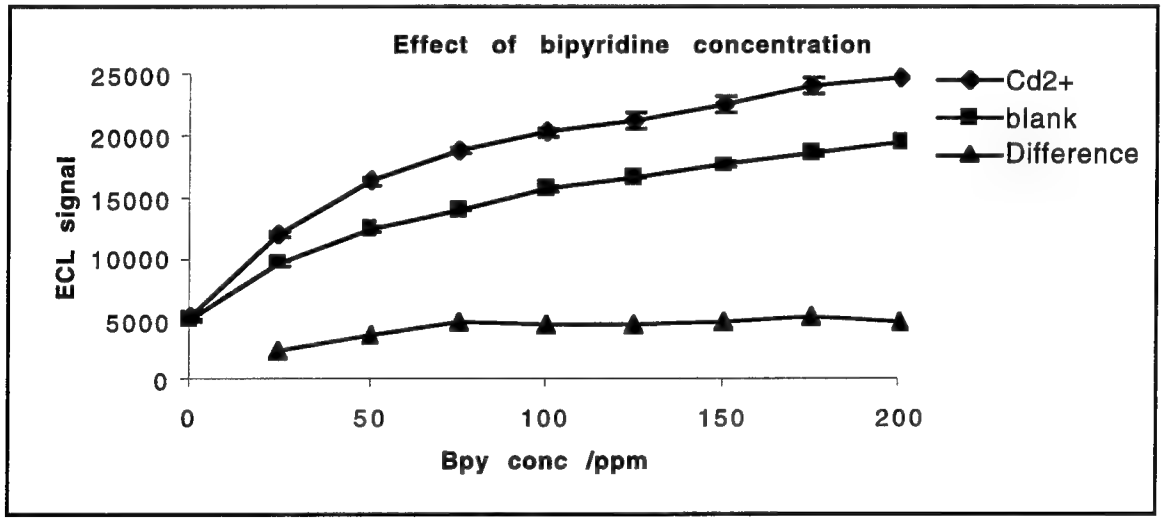
diatomic ions showed a large reduction in the Cd^{+2} control signal. Ions that caused a reduction in signal included Al (28.7% of control), Ca (26.8%), Co (2.3%), Cu (13.3%), Fe (28.8%), Ni (13.2%), Sn (1.1%), V (0.6%), Zn (4.3%). At 1 ppm (100 µl of 10 ppm standard) the smallest inhibition of the Cd signal was Co at 42.9% of the control signal.

Bipyridine

The examinations of 1,10-P and terpyridine parameters were repeated for bipyridine. As with each previous ECL measurement, two samples types were prepared. One contained cadmium and the other DI water as a blank. Again, the water blank was subtracted from the cadmium containing sample after measurement and the difference was recorded.

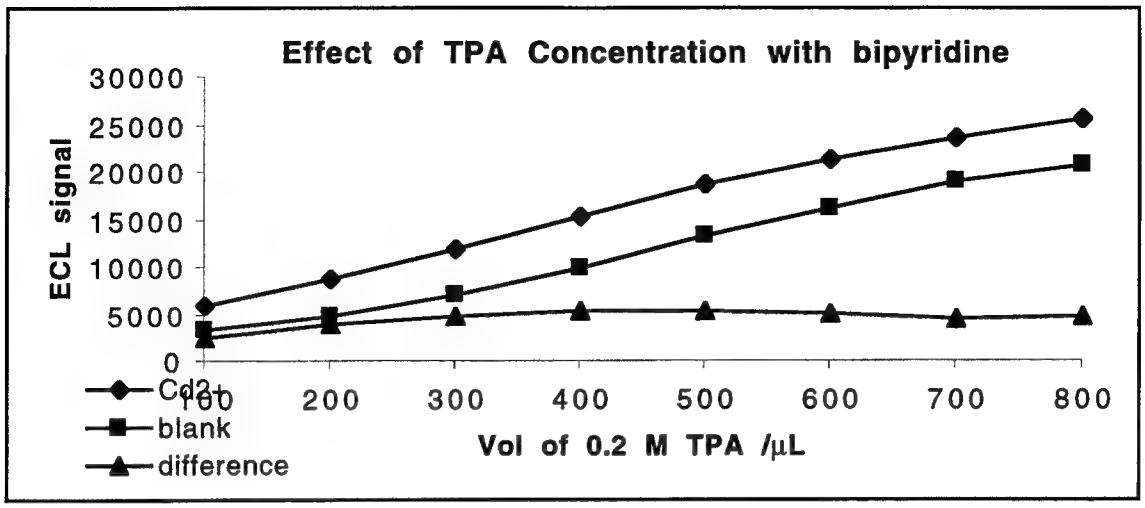
Effect of bipyridine concentration

Bipyridine concentrations from 0-200 ppm were examined using stock solution of 1 mg/ml. The ECL values obtained at bypyridine concentration of 75 ppm and greater were similar in value. This find suggested 100 µl 1mg/ml terpyridine is ideal in terms of addition simplicity using an automatic pipetting device. The final concentration of bipyridine in the 1 ml sample is 100 ppm.



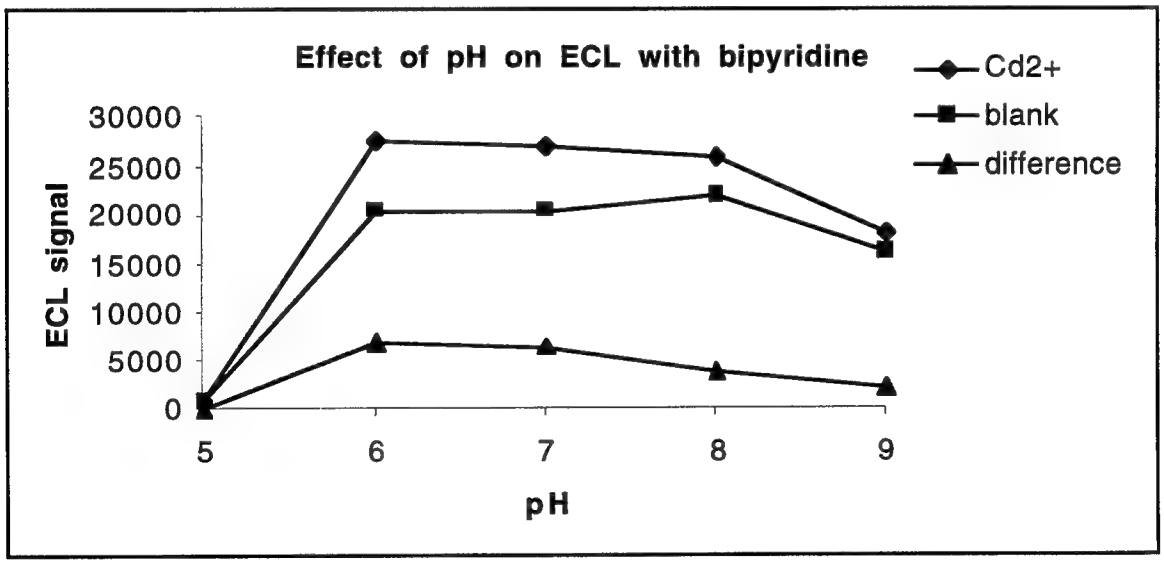
Effect of TPA Concentration

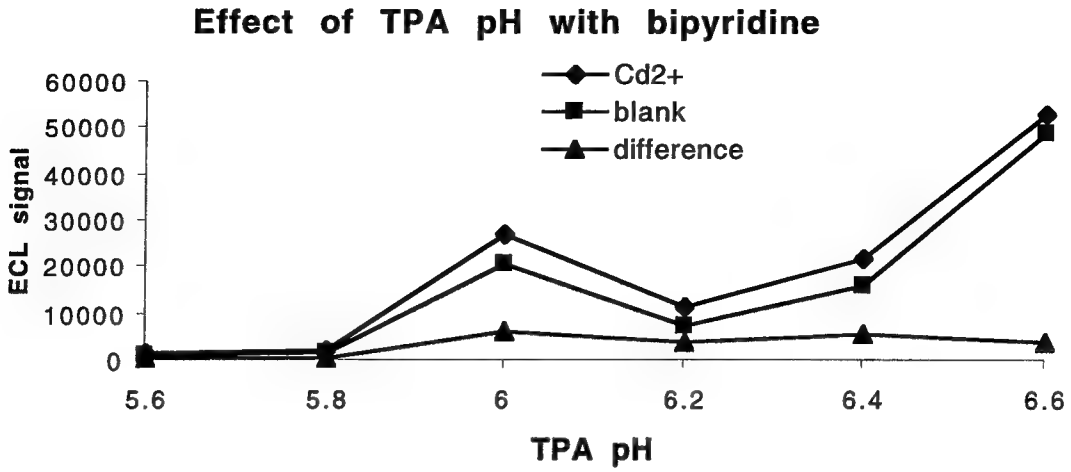
The volume of 0.2 M TPA was varied from 0 to 800 µl per each run. The ECL output was shown to be linearly increasing with increasing TPA volume. No advantage was found in excess of 400 µl per sample. The TPA addition to a sample was 400 µl 0.2 M TPA with 400 µl buffer without TPA (0.1 M TPA final concentration).



Effect of TPA pH

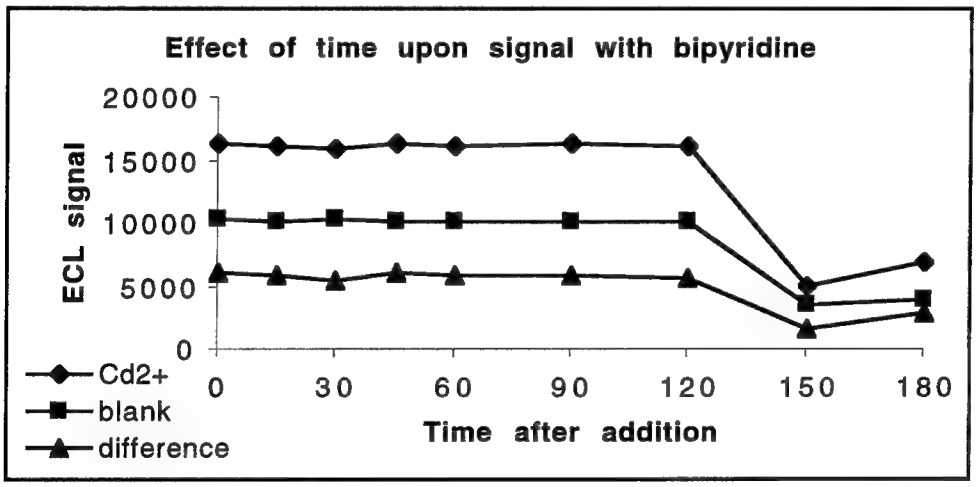
The effect of TPA pH was examined at pH values 5, 6, 7, 8, and 9 with 100 µl 10 ppm Cd^{+2}, 100 µl 1 mg/ml bipyridine, and 800 µl 0.1 M TPA. A peak in ECL was noticed between 5 and 7. Further pH values examined were 5.6, 5.8, 6.0, 6.2, 6.4, and 6.6. From the data collected, TPA at pH 6 maximized the ECL signal.





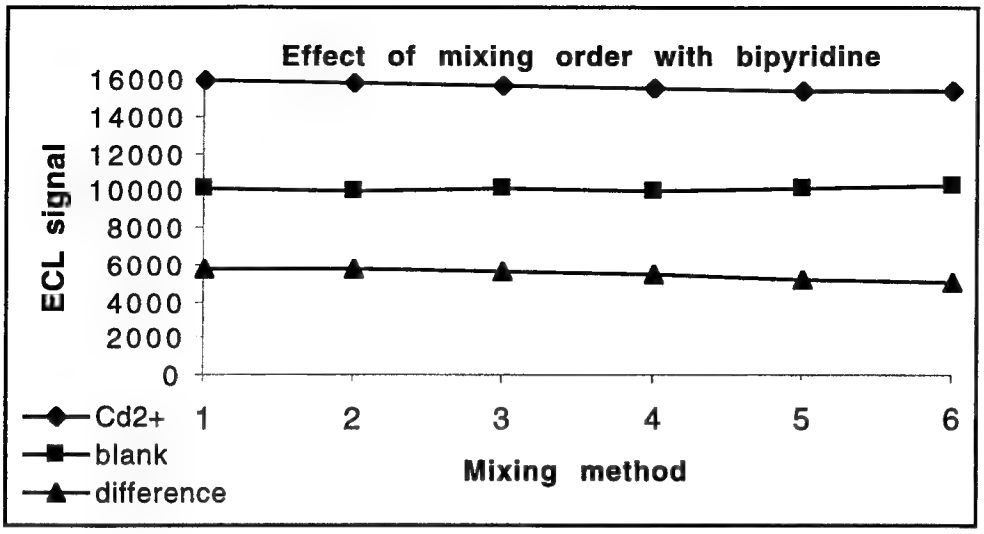
Effect of Time Upon Signal

Samples were mixed in the previously described order and allowed to interact for 0, 15, 30, 45, 60, 90, 120, 150, and 180 minutes. Allowing the reagents to mix for excess of 0-15 minutes showed little improvement on the ECL signal, until a noticeable decrease in signal at 120 minutes.



Effect of Mixing Order

The samples were examined for any ECL enhancement due to mixing order. The runs were as follows: 1)ion, bipyridine, TPA, 2) ion, TPA, bipyridine, 3) bipyridine, ion, TPA, 4) bipyridine, TPA, ion, 5) TPA, bipyridine, ion, 6) TPA, ion, bipyridine. The concentrations and amounts of reagents added were the ones reported in the examinations stated above. The output of the experiment showed mixing method 2 to be highest of those examined.



Cadmium Calibration

Due to time limitations, the cadmium calibration utilizing bipyridine parameters was not able to be performed.

Interference for Added Ions

Due to time limitations, the examination of potential interfering ions utilizing bipyridine parameters was not able to be performed.

Conclusions

In conclusion, ECL has been shown in preliminary studies to be a viable method in which an ion sensor can be based. The results of ligand optimization are summarized below. The protocols developed for ECL measurements can be applied to a sensor configuration for use *in situ* or in conjunction with analytical devices such as ion exchangers for more accurate ion detection.

Additional experiments will include: 1) remaining parameters of bipyridine calibration , 2) examination of potential ion interference, and 3) similar detailed examinations of organic candidates able to detect environmentally relevant ions in addition to cadmium. Specifically, the preliminary work will be expanded for use in a portable, in-line prototype ECL sensor discussed in "Novel Electrochemiluminescence Reactions and Instrumentation," Anthony Andrews, AFOSR, Washington, DC, Armstrong Laboratory/Environics Directorate, Tyndall AFB, FL, 1997 in which additional parameters such as electrode voltage and sample residence times can be manipulated and explored for further signal enhancement. This

sensor will first duplicate the fixed volume ion detection method as discussed in this paper and be expanded to examine in-line flow ion detection.

The advancements presented will help detect, thus better gauge, pollution levels. Applications of the technology presented include use in monitoring of pollution remediation efforts, screening and quantification of metals in metallic processes, and ultimately advance researchers in the growing demand for developing technologies that help limit metallic pollution.

Summary of parameters study.

	1, 10-phenanthroline	terpyridine	bipyridine
Effect of organic concentration	200 µl 0.1 mg/ml	100 µl 0.1 mg/ml	100µl 1 mg/ml
Effect of TPA Concentration	700µl 0.2 M	800 µl 0.2 M	400 µl 0.2 M + 400µl buffer
Effect of TPA pH	8.2	7.5	6.0
Effect of Time Upon Signal	15 minutes	15 minutes	15 minutes
Effect of Mixing Order	1,10-phenanthroline, TPA, ion	ion, TPA, terpyridine	ion, TPA, bipyridine
Cadmium Calibration	$r^2 = 0.9994$	$r^2 = 0.9922$	N/A
Interference for Added Ions at 1ppm	Co, 79.9% of Cd^{+2} control signal	Co, 42.9% of Cd^{+2} control signal	N/A

References

_1. Bruno, J.G., *et al.*, *Electrochemiluminescence from Tunicate, Tunichrome-Metal Complexes and Other Biological Samples.* J Biolumin Chemilumin, 1996. **11**: p. 193-206.

2. Bruno, J.G., S.D. Collard, and A.A. Andrews, *Further Characterization of Tunicate and Tunichrome Electrochemiluminescence.* J Biolumin and Chemilumin, 1996.

3. Jones, P., T. Williams, and L. Edbon, *Development of a Novel multi-element detection system for trace metal determination based on chemiluminescence after separation by ion chromatography.* Analytica Chimica Acta, 1990. **237**: p. 291-298.

4. Igen, I., *ORIGEN Analyzer Operator Manual*, . 1995.

Molecular Typing of Candida Parasilosis Via Amplified Fragment Length Polymorphism and Repetitive Sequence-Based PCR

James M. Tickner
Graduate Student
Institute of Molecular Biology and Medicine

University of Scranton
Scranton, PA 18510-4625

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

And

Armstrong Laboratory

September 1997

MOLECULAR TYPING OF *CANDIDA PARAPSILOSIS* VIA AMPLIFIED FRAGMENT LENGHTH POLYMORPHISM AND REPETITIVE SEQUENCE-BASED PCR

James M. Tickner

James M. Tickner

University of Scranton

Graduate Studies

Institute of Molecular Biology and Medicine

Scranton, PA 18510-4625

Abstract

Two fingerprinting techniques were used on thirty-six isolates of *Candida parapsilosis* to study their genetic relatedness. Each organism was cultured and genomic DNA was consequentially isolated. Isolation techniques were optimized at the University of Scranton, Scranton, PA and Armstrong Laboratories, Brooks AFB, San Antonio, TX. Two fingerprinting methods selected, viz. repetitive sequence-based PCR and amplified fragment length polymorphism. Repetitive sequence-based PCR results showed 33 of the 35 isolates to be identical when fragments were separated in a agarose gel. The isolates were randomly collected from across America and were expected to show non-identical fingerprints. Since the rep-PCR did not discriminate the strains, we decided to compare their results to an amplified fragment polymorphism technique (AFLP). AFLP analysis was used to visualize a large number of amplified DNA restriction fragments simultaneously. AFLP results corroborated with the rep-PCR data and placed the isolates in 3 distinct groups. These data indicate that rep-PCR is equally sensitive to AFLP and can be utilized routinely to study genetic relatedness in epidemiological studies.

MOLECULAR TYPING OF *CANDIDA PARAPSILOSIS* VIA AMPLIFIED FRAGMENT LENGHTH POLYMORPHISM AND REPETITIVE SEQUENCE-BASED PCR

James M. Tickner

Introduction

Therapeutic treatment of cancer and AIDS patients is quickly advancing and altering the hospital patient population (Rinaldi-97). These medical treatments are allowing the patients to live longer then ever. With increased longevity and decreased morbidity, this vastly growing immunocompromised population is at high risks to opportunistic pathogens (Niesters-93). *Candida parapsilosis* is one of the opportunistic deadly mycotic infections which now has acquired new prominence in the hospital environment and is a leading cause of nosocomial outbreaks across the county. It is of grave importance to maintain and expand species specific research due to this hospital-acquired infection (Merz-94). Physicians now use several antifungal medications to treat infected hosts. However, since the therapies are highly toxic, doctors are very hesitant to prescribe medication until exact genotyping is determined (Buchman-90). Techniques used today are labor-intensive, time consuming and require substantial expertise. Due to this slow process species determination is often done during an autopsy which needs to be remedied. Thus, a rapid, specific and cost effective diagnostic typing tool needs to be developed which can be routinely used by technicians in a hospital setting (Morace-96, Redkar-96, DelVecchio-95, Pfaller-92).

MOLECULAR TYPING OF *CANDIDA PARAPSILOSIS* VIA AMPLIFIED FRAGMENT LENGHTH POLYMORPHISM AND REPETITIVE SEQUENCE-BASED PCR

James M. Tickner

Materials and Methods

Yeast Strains. Thirty-six isolates of *C. parapsilosis* were used in this study. Thirty-four of these strains were collected by Dr. Mike Rinaldi, University of Texas Health Science Center, San Antonio, Texas, (97-446 through 97-1035). Strains Scr-A and Scr-B were obtained from the American Type Culture Collection, Rockville, MD. All isolates were cultured and grown at room temperature on a 2% YPD agar medium (Difco Laboratories, Detroit Mich.).

DNA isolation. Single colonies were inoculated in 2.5 ml of YPD broth containing 1% yeast extract, 2% Bacto Peptone, and 2% glucose (Difco Laboratories)and inoculated overnight with shaking at 175 rpm at 30°C in falcon tubes. A 25% glycerol stock was prepared from 1.0 ml of the above growth and stored at -80°C. DNA from each strain was isolated with methods described by Scherer and Stevens (Scherer-87) with slight modifications. In summary, the yeast cells were collected by centrifugation (1.5 ml inoculated broth at 3,000 rpm for 1 min), suspended in 1.0 ml of 1 M Sorbitol, centrifuged again, and resuspended in 1 ml of the following mix (4ml of 1 M sorbitol-50 mM potassium phosphate buffer (pH 7.5), plus 4 µl B-mercaptoethanol and 15 µl yeast lytic enzyme (ICN Biomedicals, Aurora, Ohio)). After a 30 minute incubation at 30°C shaking at 35 rpm, spheroplast were pelleted and resuspended in 0.5 ml of the following mix (250 ul of 500 mM EDTA (pH 7.5) with 50 µl of 10% sodium dodecyl sulphate with 2.2 ml distilled water addition). To this, 0.3 µl proteinase K , 20 mg/ml, (Amresco, Solon, Ohio) was added followed by 2-3 hour incubation at 50°C. The sample was extracted three times with equal volumes of phenol/chloroform (1:1; vol/vol) by use of phase lock gel tubes and then extracted with a 1:1 volume of chloroform isoamyl alcohol mixture. After precipitation with 2 volumes 100% ethanol , centrifugation at top speed for 5 minutes, and drying, the DNA was resuspended in 100 ul TE buffer followed by a 3 ul

RNase addition. This was incubated at 37°C for 30 minutes. DNA was then precipitated by isopropyl (4 vols) and 70% ethanol (1 ml), allowed to air dry, and resuspended in 100 ul distilled water. After determination of the concentration with the GeneQuant RNA/DNA Calculator (Pharmacia Biotech, Piscataway, N.J.), about 700 ng of DNA in combination with 5 ul loading dye was run on a 0.7% (wt/vol) SeaKem GTG aqarose gel (FMC BioProducts, Rockland, ME) to determine the quality of DNA.

Rep-PCR. The PCR was performed in 50 µl standard PCR tubes containing 200 ng DNA, 50 pmol each primer Care-21 and Care-22, or 100 pmol Com-21 if used alone, 1.25 U of *Taq* polymerase (Perkin-Elmer, Branchburg, N. J.), 0.2 mM each deoxribnucleotide triphosphate, 1X concentration buffer. The mixture was subjected to an initial denaturation of 94°C for 5 min, followed by 40 cycles of denaturation at 94°C for 1 min, primer annealing at 45°C for 1 min, and extension at 72°C for 2 min. A final extension at 72°C was also performed.

Gel Electrphoresis. All amplicons from the various yeast isolates were analyzed on 2.0% (wt/vol) SeaKem GTG agarose gels (FMC BioProducts, Rockland, ME) containing ethidium bromide (0.5 µg/ml) made in a 0.5X concentration of Trisacetate-EDTA buffer. A 100-bp DNA ladder was run concurrently in each gel to determine molecular sizes of amplicons. The gels were run at approximately 87v for 2.5 hrs. The gels were photographed using Polaroid 667 film. The banding patterns were compaired visually and strains were placed in different groups.

Amplified Fragment Length Polymorphism technique was performed as procedures from AFLP Microbial Fingerprinting protocol (PE Applied Biosystems, Foster City, California). In short, the technique involves three steps: (1) restriction of genomic DNA using *MseI* and *EcoRI* resriction enzymes followed by a ligation of *MseI* and *EcoRI* adaptors

provided with the AFLP kit, (2) Preselective amplication of a subset of the restricted DNA using adaptor based primers and (3)selective amplification with specially selected primers using touchdown-PCR. This was followed by analysis of fingerprints on a ABI 371A machine with Genescan software.

Results

Analysis of rep-PCR results and typing of *C. Parapsilosis*. Four primer combinations (CARE-21 plus CARE-22, COM-21 alone, CARE-22 plus COM-21, CARE-21 plus COM-21) were first utilized to evaluate which combinations worked best on *C. parapsilosis*. Generation of 6-10 distinct bands of equal intensities was used as a criteria for selecting primers. Ultimately the primer combination of CARE-21 plus CARE-22 was selected for typing of 34 isolates. Only two strains (97-460, 97-828) demonstrated unique fingerprints, which were collected from the states of Texas and Florida respectively and were different from each other as well. All 32 other isolates showed identical fingerprints on the 2% SeaKem GTG agarose gel [Fig. 1A and 1B]. Based on these patterns isolates were placed in three distinct groups, viz., A (32), B (1), and C(1). The groups to which individual isolates belong is shown in Table 1 with the state in which it was collected.

Analysis of AFLP Results. All 34 isolates of the *C. parapsilosis* were employed to demonstrate the application of this technique. DNA was previously isolated for use in rep-PCR studies and consequentially used in the AFLP analysis. Thirty-two of the *C. parapsilosis* strains showed identical fingerprinting patterns while the two remaining isolates (97-460, 97-828) were unique to each other as well as the 32 others [Fig. 2].

FIGURE 1A

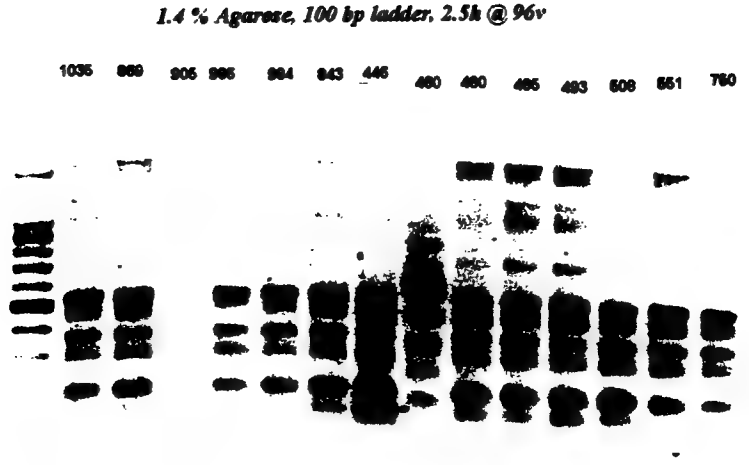
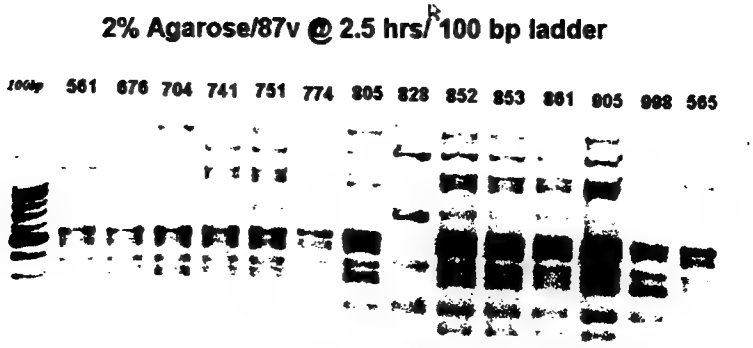


FIGURE 1B



Gels Fig. 1A & 1B. Rep-PCR fingerprints of *C. parapsilosis* isolates. Primers Ca-21 and Ca-22 were used in PCRs, and the amplicons were resolved on a 2% SeaKem GTG agarose gel (for details, refer to Materials and Methods. Molecular size markers (100 bp) are in left lanes.

James M. Tickner

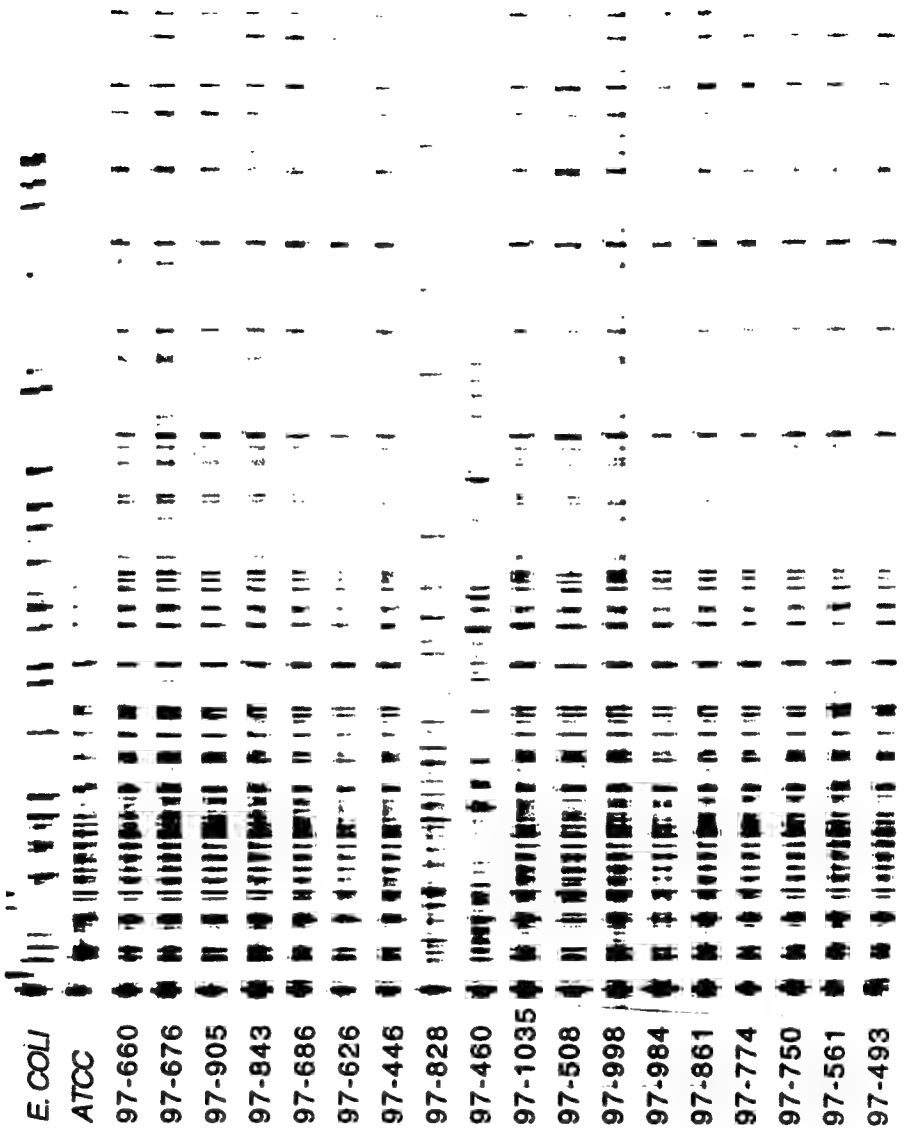
TABLE 1. *C. parapsilosis* isolates

Isolate Number	State Collected	Assigned Type
97-1035	PA	C
97-998	GA	C
97-986	PA	C
97-984	CO	C
97-905	PA	C
97-869	TX	C
97-861	CA	C
97-853	CA	C
97-852	AZ	C
97-843	NH	C
97-828	FL	B
97-805	TX	C
97-774	TX	C
97-751	MA	C
97-750	MO	C
97-741	TN	C
97-686	MO	C
97-704	MI	C
97-676	PA	C
97-662	CA	C
97-660	TX	C
97-655	FL	C
97-626	IL	C
97-612	WA	C
97-565	MI	C
97-561	TN	C
97-556	MA	C
97-551	RI	C
97-460	TX	A
97-508	FL	C
97-493	LA	C
97-485	CA	C
97-480	ME	C
97-446	PA	C
SCR-A*	ATCC	C

* obtained from the American Type culture collection ATCC 22019

MOLECULAR TYPING OF *CANDIDA PARAPSILOSIS* VIA AMPLIFIED FRAGMENT LENGHTH POLYMORPHISM AND REPETITIVE SEQUENCE-BASED PCR

James M. Tickner



AFLP Figure AFLP fingerprints of 19 representative samples of *C. parapsilosis* strains with E. Coli control strain in left lane.

Discussion

The objective of this study was to use a rep-PCR based fingerprinting technique to discriminate among different isolates of *C. parapsilosis*. Rep-PCR technique was selected since it is informative, reliable, reproducible, fast and cost effective. The DNA fingerprinting patterns by rep-PCR were found to be identical with the exception of two isolates. This was least expected since all isolates were collected from non-common sources across the country. These data indicates conservation in genotype of all pathogenic *C. parasilosis* which, in turn, could be used in effective treatment of these infections.

In order to substantiate rep-PCR results, we utilized AFLP analysis on the same samples. AFLP approach is a recent, novel and powerful DNA fingerprinting technique which can provide greater sensitivity in relatedness among pathogenic isolates. That the results of AFLP upon Genescan analysis confirmed the rep-PCR fingerprinting results is an important observation. This not only confirms the sensitivity of rep-PCR, but indicates total conservation of pathogenic genotype across the country.

It is also possible that the same strain of *C. parapsilosis* is causing the infections that are seemingly becoming more common in the hospital setting. This is an intriguing observation since *E. Coli* 0157:H7 isolated from worldwide outbreaks also show identical fingerprinting patterns by rep-PCR.

It is clearly important to see if our findings hold up for more strains of *C. parapsilosis* isolated from different sources. Additional isolates are currently being examined by our laboratory. It would be interesting to compare our results with karyotyping techniques from pulsed field gel analysis which is the ultimate technique for DNA fingerprinting. In

conclusion, we have used rep-PCR to type 34 isolates of *C. parapsilosis* and have shown that 32 of these leave identical genotype. Interestingly, this data was supported by AFLP technique.

References

Bart-Delabesse, E., H. V. Deventer, W. Goessens, J. Poirot, N. Lioret, A. V. Belkum, and F. Dromer. 1995. Contribution of molecular typing methods and antifungal susceptibility testing to the study of a candidemia cluster in a burn care unit. J. Clin. Microbiol. 33:3278-3283.

Bingen, E. H., E. Denamur, and J. Elion. 1994. Use of ribotyping in epidemiological surveillance of nosocomial outbreaks. Clin. Microbiol. Rev. 7:311-327.

Brosius, J., A. Ullrich, M. A. Raker, A. Gray, T. J. Dull, R. R. Gutell, and H. F. Noller. 1981. Construction and fine mapping of recombinant plasmids containing the rrnB ribosomal RNA operon of *E. coli*. Plasmid 6:112-118.

Bushman, T. G., M. Rossier, W. G. Merz, and P. Charache. 1990. Detection of surgical pathogens by in vitro DNA amplification of *Candida albicans* by in vitro amplification of a fungus-specific gene. Surgery 108:338-347.

Carruba, G., E. Pontieri, F. DeBernardis, P. Martino, and A. Cassone. 1991. DNA Fingerprinting and Electrophoretic Karyotype of Environmental and Clinical Isolates of *Candida parapsilosis*. J. Clin. Microbiol. 29:916-922.

DelVecchio, V. G., J. M. Petroziello, M. J. Gress, F. K. McCleskey, G. P. Melcher, H. K. Crouch, and J. R. Lupski. 1995. Molecular genotyping of methicillin-resistant *Staphylococcus aureus* via fluorophore-enhanced repetitive-sequence PCR. J. Clin. Microbiol. 33:2141-2144.

Deresinski, S. C., K. V. Clemons, C. A. Kemper, K. Roesch, B. Walton, and D. A. Stevens. 1995. Genotypic Analysis of Pseudoepidemic Due to Contamination of Hanks' Balanced Salt Solution with *Candida parapsilosis*. J. Clin. Microbiol. 33:2224-2226.

Diming, L., L. Wu, M. G. Rinaldi, and P. F. Lehman. 1995. Three distinct Genotypes within *Candida parapsilosis* from Clinical Sources. J. Clin. Microbiol. 33:1815-1821.

Edelstein, P. H. 1981. Improved semiselective medium for isolation *of Legionella pneumophila* from contaminated clinical and environmental specimens. J. Clin. Microbiol. 14:298-303.

Ferdows, M. S. and A. G. Barbour. 1989. Megabased-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. Proc. Natl. Acad. Sci. 86:5969-5973.

Gaia, V., C. Poloni, and R. Peduzzi. 1994. Epidemiological typing of *Legioella pneumophila* with ribotyping: report of two clinical cases. Eur. J. Epidemiol. 10:303-306.

Grimont, F., M. Lefevre, E. Ageron, and P. A. D. Grimont. 1989. rRNA gene restriction patterns of *Legionella* species: a molecular identification system. Res. Microbiol. 140:615-626.

Hellstein, J., H. Vawter-Hugart, P. Fotos, J. Schmid, and D. R. Soll. 1993. Genetic similarity and phenotypic diversity of commensal and pathogenic strains of *Candida albicans* isolated from the oral cavity. J. Clin. Microbiol. 31:3190-3199.

Klein, J. J., and C. Watanakunakorn. 1979. Hostpital-acquired fungemia: its natural course and clinical significance. Am. J. Med. 67:51-58.

Merz, W. G 1994. Clinical mycology laboratory: meeting the challenge of the 1990s. Infect. Dis. Clin. Prac. 3(Suppl 2):S60-S67.

Morace, G., M. Sanguinetti, B. Posteraro, G. L. Cascio, and G. Fadda. 1997. Identification of various medically important *Candida* species in clinical specimens by PCR-restriction enzyme analysis. J. Clin. Microbiol. 35:667-672.

Niesters H. G. M., W. H. F. Goessens, J. F. M. G. Meis, and W. G. V. Quint. 1993. Rapid, polymerase chain reaction-based identification assays for *Candida* species. J. Clin. Microbiol. 31:904-910.

Picardeau, M., G. Prod'hom, L. Raskine, M. P. LePennec, and V. Vincent. 1997. Genotypic characterization of five subspecies of *Mycobacterium kansasii.* J. Clin. Microbiol. 35:25-32.

Pfaller, M. A. 1992. Epidemiological typing methods for mycoses. Clin. Infect. Dis. 14(suppl 1):S4-S10.

Prasad, J. K., I. Feller, and P. D. Thomson. 1987. A ten-year review of *Candida* sepsis and mortality in burn patients. Surgery 101:213-216.

Redkar, R. J., M. P. Dube, F. K. McCleskey, M. G. Rinaldi, and V. G. DelVecchio. 1996. DNA fingerprinting of *Candida rugosa* via repetitive sequence-based PCR. J. Clin. Microbiol. 34:1677-1681.

Reinhardt, J. F., P. J. Ruane, L. J. Walker, and W. L. George. 1985. Intravenous catheter-associated fungemia due to *Candida rugosa.* J. Clin. Microbiol. 22:1056-1057.

Replogle, J., W. D. Lord, B. Budowle, T. L. Meinking, and D. Taplin. 1994. Identification of host DNA by amplified fragment length polymorphism analysis: preliminary analysis of human crab louse (Anoplura: Pediculidae) excreta. J. Med. Entomol. 31:686-690.

Rinaldi, M. 1997. Fungi can be friends or foes...and they're getting smarter. The News: The University of Texas Health Science Center at San Antonio Newspaper: vol. XXX, no. 29, July 18, 1997.

Sandhu, G. S., B. C. Kline, L. Stockman, and G. D. Roberts. 1995. Molecular probes for diagnosis of fungal infections. J. Clin. Microbiol. 33:2913-2919.

Scherer, S., and D. A. Stevens. 1987. Application of DNA typing methods to epidemiology and taxonomy of *Candida* species. J. Clin. Microbiol. 25:675-679.

Scherer, S., and D. A. Stevens. 1988. A *Candida albicans* dispersed, repeated gene family and its epidemiologic applications. Proc. Natl. Acad. Sci. 85:1452-1456.

Schmid, J., E. Voss, and D. R. Soll. 1990. Computer-assisted methods for assessing strain relatedness in *Candida albicans* by fingerprinting with moderately repetitive sequence Ca3. J. Clin. Microbiol. 28:1236-1243.

Sugar, A. M., and D. A. Stevens. 1985. *Candida rugosa* in immunocompromised infections: case reports, drug susceptibility and review of the literature. Cancer 56:318-320.

Taylor, N. S., J. G. Fox, N. S. Akopyants, D. E. Berg, N. Thompson, B. Shames, L. Yan, E. Fontham, F. Janney, F. M. Hunter, and P Correa. 1995. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA fingerprinting. J. Clin. Microbiol. 33:918-923.

Valsangiacomo, C., F. Baggi, V. Gaia, T. Balmelli, R Peduzzi, and J. C. Piffaretti. 1995. Use of amplified fragment length polymorphism in molecular typing of *Legionella pneumophila* and application to epidemiological studies. J. Clin. Microbiol. 33:1716-1719.

Vaneechoutte, M. 1996. Review: DNA fingerprinting techniques for microorganisms: a proposal for classification and nomenclature. Mol. Biotech. 6:115-141.

Wildfeuer, A., R. Schlenk, and W. Friedrich. 1996. Detection of *Candida albicans* DNA with yeast-specific primer system by polymerase chain reaction. Mycoses 39:341-346.

A TEST OF THE MISATTRIBUTED-ACTIVATION HYPOTHESIS OF THE REVELATION EFFECT IN MEMORY

Deanne L. Westerman
Graduate Student
Department of Psychology

Case Western Reserve University
10900 Euclid Avenue
Cleveland, OH 44106-7123

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1997

A TEST OF THE MISATTRIBUTED-ACTIVATION HYPOTHESIS OF THE REVELATION EFFECT IN MEMORY

Deanne L. Westerman
Ph.D. student

Department of Psychology
Case Western Reserve University
Cleveland, Ohio 44106-7123

Abstract

The revelation effect refers to the tendency to call an item on a recognition test "old" if a cognitive task that involves processing of a similar stimulus precedes it (Watkins & Peynircioglu, 1990). This effect occurs for both old and new test items and results in a pattern of higher hits and false alarms for items in the revelation condition. It has been proposed that the revelation effect reflects enhanced familiarity for test items in the revelation condition (Luo, 1993; Westerman & Greene, in press). A satisfactory account of the revelation effect has been elusive. The present experiment was designed to test one potential explanation for this effect, which proposes that the revelation effect reflects an increase in the familiarity of the test item due to the activation of other list items during the preceding cognitive task.

A TEST OF THE MISATTRIBUTED-ACTIVATION HYPOTHESIS OF THE REVELATION EFFECT IN MEMORY

Recent interest in so-called memory illusions has led to a growing body of research that is aimed toward understanding the principles that lead a person to falsely recall or recognize an item or an event (e.g., Payne, Elie, Blackwell & Neuschatz, 1996; Reinitz, Lammers, & Cochran, 1992; Roediger, 1996; Roediger & McDermott, 1995). Among this research are several studies that show that illusions of memory can occur due to conditions that are present during a recognition test (e.g., Jacoby & Whitehouse, 1989; Lindsay & Kelly, 1996; Whittlesea, 1993; Whittlesea, Jacoby, & Girard, 1990). One example of recognition induced by specific test conditions is the revelation effect in memory. The "revelation effect" is the term used by Watkins & Peynircioglu (1990) to describe their finding that recognition test items that appear in a distorted form (e.g., as an anagram or as a word fragment) and are solved or "revealed" before the recognition decision are more likely to be called "old" than items that are presented intact on the recognition test. This tendency to claim that revealed items are old occurs for both old and new test items, but the magnitude of the effect is often significantly greater for new items (LeCompte, 1995, Experiments 1 and 2; Watkins & Peynircioglu, 1993, Experiment 3; Westerman & Greene, 1996, Experiments 2, 5, and 7; Westerman & Greene, in press; Experiment 4). As a result, recognition performance tends to be less accurate in the revelation condition because the increase in false alarms is usually greater than the increase in hits.

Surprisingly, a revelation effect has been found in many situations that do not involve the revelation of a test item. Westerman and Greene (1996) found a revelation

effect when the word that was revealed was not the word that was subsequently recognized, but rather a different word that had not been presented on the study list and that bore no obvious relationship to the test word. Recent research has shown that the term, "revelation effect" appears to be a misnomer, as a revelation effect has been found using a variety of tasks. For instance, tasks that involved counting letters, generating synonyms, rearranging the letters of illegal nonwords, and recalling a short list of letters, all led to a revelation effect. That is, test items that followed these tasks were more likely to be called old than test items that did not follow these tasks (Westerman & Greene, in press).

The robustness and generality of the revelation effect have been well established. A revelation effect has been found in numerous experiments using a variety of cognitive tasks in the revelation phase (Watkins & Peynircioglu, 1990; Westerman and Greene, in press). The effect has been found in both within and between-subjects designs (Westerman & Greene, 1996), and when words, numbers (Watkins & Peynircioglu, 1990) and pictures (Luo, 1993) are used as stimuli. The revelation effect does not depend on participants' successful completion of the interpolated task; a revelation effect occurs even when participants can solve few of the revelation problems (Westerman & Greene, in press). A revelation effect has been reported on several types of episodic memory tests; revelation has been found to affect frequency judgments, list discrimination (Westerman & Greene, 1996), and responses on tasks that separate the contribution of familiarity and recollective experience (LeCompte, 1995). Indeed, the only memory tests that have been shown to be unaffected by revelation are those that require judgments that are not based on episodic memory. When no study list was presented, revelation was not found to influence

judgments of category typicality, lexicality, or word frequency (Watkins & Peynircioglu).

One factor that has been found to be crucial to the emergence of a revelation effect is the compatibility between the interpolated stimuli and the test stimuli. Using words as the study and test stimuli, a revelation effect was not found when arithmetic problems were used as the revelation task (Westerman & Greene, in press), despite prior research that showed that arithmetic problems produce a revelation effect when numbers are used as study and test items (Watkins & Peynircioglu, 1990). Additional evidence for the importance of the compatibility between the revelation and to-be recognized stimuli was found in an experiment in which a revelation effect was not found on a word recognition task when digits were used as stimuli in a memory span task, in spite of an earlier finding that found a revelation effect when a letter-span task was used (Westerman & Greene, in press).

Several accounts of the revelation effect have been proposed. One potential explanation for the pattern of higher hits and false alarms in the revelation condition is that the revelation task causes a shift in participants' decision making criteria on the recognition test. For instance, it is possible that the revelation manipulation primes the test item, which creates a bias to respond positively to it on the recognition test trial, thereby creating a pattern of higher hits and false alarms on a recognition test. (see Mandler, Nakamura, & Van Zandt, 1987 and Ratcliff & McKoon, 1995 for examples of positive response biases that result from priming). There is evidence against a positive response bias explanation of the revelation effect, however. Westerman and Greene (1996) found that revelation caused an increase in negative responses when the memory test required participants to respond Ayes to the least familiar items. In one experiment, participants

were given a list discrimination test in which they were asked to respond No to the most recently presented items. In this case revealed items received significantly more no responses than intact items. In another experiment, participants were given a recognition test in which they were asked to respond "no" to old items and to respond "yes" to new items. Similar to the results of the list-discrimination task, participants were more likely to respond no revealed items than to intact items. These experiments rule out the possibility that the revelation effect is merely a bias to respond positively to a recognition test item. However, it is possible that the revelation effect is caused by a more complex recognition bias. For instance, it could be that the interpolated task causes participants to call a test item "old" at a lower familiarity level. However, the reason why the revelation task would have such an effect on participants' response criteria is unclear.

A second explanation for the pattern of higher hits and false alarms in revelation effect experiments is that the task performed during the revelation phase increases the familiarity of a test item that follows it. One possibility that has been suggested by several authors (Luo, 1993; Peynircioglu & Tekcan, 1993) is that the revelation effect may occur due to a misattribution of familiarity (c.f., Jacoby & Whitehouse, 1989). According to this account, the processing of a stimulus during the revelation phase increases its familiarity, and the enhanced familiarity is misattributed to the test item's presence on the study list. Although this account can explain the revelation effect when the items that are processed during the revelation phase match those items that are recognized, it can not account for a revelation effect when the stimuli that are processed do not match the words that are to-be recognized (Westerman & Greene, 1996, Westerman & Greene, in press).

An alternate account of the revelation effect, which is similar to the familiarity misattribution account, insofar as it attributes the revelation effect to enhanced familiarity of the test items in the revelation condition, has been offered by Westerman and Greene (in press). This account can be described in terms of the global-matching approach to recognition memory (e.g., Gillund & Shiffrin, 1994; Hintzman, 1998; Murdock, 1993; for a review, see Clark & Gronlund, 1996). According to global-matching models, the presentation of an item on a recognition test causes a number of representations in memory to become activated, with the level of activation of each representations a function of its similarity to the test item. Recognition responses are determined by the sum of the activation of all memory representations, with higher levels of activation increasing the probability of an "old" response. To explain the revelation effect, Westerman and Greene proposed that the processing of a stimulus in the revelation task leads to the activation of some memory representations, and that this activation persists for some time after the task is completed. When the recognition test item appears, the activation produced by the revelation task contributes to the activation level upon which the recognition decision is made. Because the total level of activation is higher on test trials that follow the revelation task, participants tend to respond old more often to these items than to items that do not follow the revelation task.

Westerman and Greene's (in press) account proposes that the revelation task enhances the familiarity of test items that follow it, which is caused by the activation of items from the study list. If this account is true, then the magnitude of the effect should depend on the similarity of the stimuli in the interpolated task to the items on the study list. If the stimuli processed during the interpolated task were among those on the study

list, then the revelation effect should be larger compared to a situation in which the word processed during the interpolated task was a word that had not been presented during the study phase. A larger revelation effect should be en the revealed words match some of the words from the study list, compared to a situation in which the revealed words do not match the studied words. This present study investigates this situation. In this experiment, a word fragment completion task was used as the revelation task and preceded half of the test items. Half of the words in the revelation phase had been present on the study list, and the other half were new.

Method

Participants. Participants were 109 United States Air Force recruits.

Materials. The stimuli were 210 6 to 8-letter words that were chosen from a pool supplied by Gibson and Watkins (1988). Each participant received a unique assignment of words to condition (intact or revealed) and to test status (target or lure), as well as a unique ordering of words in both the study and test phases. Each study list consisted of 100 targets and 10 untested buffers. The buffers occupied the first five and last five serial positions on the study list and were the same for each participant. Eighty words served as lures on the recognition test. Twenty new words were used for the revelation phase.

Procedure. Participants were tested in a large group on individual computers. Participants were told that they would see a list of words that they should try to remember. Each word was presented individually at a 2 s per item rate. Immediately following the study list, participants were given instructions for the revelation/ recognition phase. The recognition test consisted of 80 targets and 80 lures randomly intermixed and presented individually to the participants. For the recognition test trials, participants saw

the test word on the computer screen with instructions to click a box labeled "yes" if the word was on the study list and box labeled "no" if the word was not on the study list. Half of the test trials were preceded by a word fragment completion task (revelation condition); the other half were not (control condition). For half of the word fragments the solution was a word that had been presented on the study list; for the other half of the word fragments, the solution was a new word which had not been presented on the study list.

For the word fragment completion task, participants saw a word that had all but one letter missing (e.g., R-------). The solution to the word fragment was gradually revealed at a 1 letter per second rate. Participants were instructed to try to solve the word fragment with as few letters present as possible. Participants were told to press the "enter" key as soon as they knew the answer to the word fragment. When participants pressed the enter key, the solution to the word fragment stopped being revealed and they received a prompt to type in the solution to the fragment. After the entire word was revealed, or the word fragment was solved, the recognition test word appeared. Participants indicated whether the test word had been on the list by clicking "yes" or "no". Target words that were presented in the word fragment completion phase were not tested on the recognition test.

Results

The results (expressed as proportions of "old" responses) are summarized in Table 1. Replicating previous findings (Westerman and Greene, 1996; Westerman & Greene, in press), a revelation effect occurred in spite of the mismatch between the revealed word and the test word. A 2 X 2 (old/new status by recognition condition) repeated measures analysis of variance (ANOVA) was conducted on the percentage of items given "old"

responses. In all analyses reported here, a .05 significance criterion was used. As expected, there was a main effect for the items' actual status; old items were more likely to be classified as old than new items, $F(1, 108) = 343.77$, $MSE = 281.37$. More importantly, words that were recognized immediately after working on a word fragment were more likely to be called old than words that were not preceded by this task, $F(1, 108) = 108.24$, $MSE = 152.61$. There was also a significant interaction between the status of the test item and it's revelation condition, $F(1, 108) = 11.15$. Replicating previous finding, the magnitude of the revelation effect was larger for new items than old items; that is, revelation caused a larger increase in false alarms than in hits. No support was found for the hypothesis that a larger revelation effect would be obtained when the revealed item was a non-tested item from the study list. A revelation effect occurred for both study-list fragments and new fragments, and the magnitude of the revelation effect was virtually identical for both old items $t(108) = 0.26$, $p > .75$,

and new items, $t(108) = 0.69$, $p > .45$.

Discussion

A robust illusion of recognition (a revelation effect)occurred in this experiment. Test items that followed the word-fragment completion task were more likely to be attributed to the study list than test items that did not follow this task. This was true for both old and new items, however the magnitude of the illusion was larger for new items. This resulted in a larger increase in the percentage of false alarms than an increase in hits. Therefore, accuracy was reduced in the revelation condition.

A satisfactory explanation for the revelation effect remains elusive. One account, which proposes that the revelation effect is caused by an activation of items from the study

list during the revelation task, was not supported. In this experiment the magnitude of the revelation was nearly identical when the word that was revealed was a non-tested item from the study list and when the word that was revealed was a new item that had not been presented on the study list.

References

Gibson, J. M. & Watkins, M. J. (1988). A pool of 1,086 words with unique two-letter fragments. Behavior Research Methods, Instruments & Computers, 20, 390-397.

Gillund, G. & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-67.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace model. Psychological Review, 95, 528-551.

Jacoby, L. L. & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. Journal of Experimental Psychology: General, 118, 126 -135.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language, 30, 513 - 541.

LeCompte, D. C. (1995). Recollective experience in the revelation effect: Separating the contributions of recollection and familiarity. Memory & Cognition, 23, 324 - 334.

Lindsey, S. D. & Kelly, C. M. (1996). Creating illusions of familiarity in a cued recall remember/know paradigm. Journal of Memory and Language, 35, 197 - 211.

Luo, C. R. (1993). Enhanced feeling of recognition: Effects of identifying and manipulating test items on recognition memory. Journal of Experimental Psychology: Learning, Memory & Cognition, 19, 405 -413.

Mandler, G., Nakamura, Y., & Van Zandt, B. J. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. Journal of Experimental Psychology: Learning, Memory & Cognition, 13, 646 - 648.

Murdock, B. B. (1993). Todam2: A model for the storage and retrieval of item, associative and serial-order information. Psychological Review, 100, 183-203.

Payne, D. G., Elie, C. J., Blackwell, J. M. & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing and recollecting events that never occurred. Journal of Memory and Language, 35, 261 -285.

Peynircioglu, Z. F. & Tekcan, A. I. (1993). Revelation effect: Effort or priming does not create the sense of familiarity. Journal of Experimental Psychology: Learning, Memory & Cognition, 19, 382 - 388.

Reinitz, M. T., Lammers, W.J. & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. Memory & Cognition, 20 , 1 -11.

Ratcliff, R. & McKoon, G. (1995). Bias in the priming of object decisions. Journal of Experimental Psychology: Learning, Memory & Cognition, 21, 754 - 767.

Roediger, H. L.(1996). Memory illusions. Journal of Memory and Language, 35, 76 - 100.

Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory & Cognition, 21, 803 - 814.

Watkins, M. J. & Peynircioglu, Z. F. (1990). The revelation effect: When disguising test items induces recognition. Journal of Experimental Psychology: Learning,

Memory & Cognition, 16, 1012 - 1020.

Westerman, D. L. & Greene, R. L. (1996). On the generality of the revelation effect. Journal of Experimental Psychology: Learning, Memory & Cognition, 22, 1147 - 1153.

Westerman, D. L., & Greene, R. L. (in press). The revelation that the revelation is not due to revelation. Journal of Experimental Psychology: Learning, Memory, & Cognition.

Whittlesea, B. W. A. (1993). Illusions of familiarity. Journal of Experimental Psychology: Learning Memory & Cognition, 19, 1235 - 1253.

Whittlesea, B. W. A., Jacoby, L. L. & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional bias for feelings of familiarity and perceptual quality. Journal of Memory and Language, 29, 716 - 732.

Table 1.

Proportions of Old Responses by Test Item Status and Recognition Condition for Experiments 1 and 2.

	Status of Test Items	
Recognition Condition	Target (Hits)	Lure (False Alarms)
Revelation Condition		
Studied Fragment	.76	.51
Non-studied Fragment	.76	.49
Control	.67	.34